

Διαφορικός Λογισμός πολλών μεταβλητών

Πρόχειρες σημειώσεις

Μιχάλης Παπαδημητράκης

Τμήμα Μαθηματικών

Πανεπιστήμιο Κρήτης

1η εβδομάδα.

Τα στοιχεία του $\mathbb{R}^n$ είναι όλα τα n -διάστατα διανύσματα ή, ισοδύναμα, όλες οι n -άδες πραγματικών αριθμών:

$$\mathbf{x} = (x_1, \dots, x_n).$$

(Εδώ θα χρησιμοποιούμε "παχιά" γραμματοσειρά για το σύμβολο του διανύσματος $\mathbf{x}$ ενώ στον πίνακα στις διαλέξεις θα χρησιμοποιούμε το σύμβολο $\vec{x}$ με το βέλος.)

Για τον $\mathbb{R}^2$ χρησιμοποιούμε συνήθως δύο σύμβολα

$$\mathbf{x} = (x_1, x_2) \quad \text{ή} \quad \mathbf{u} = (x, y).$$

Για τον $\mathbb{R}^3$:

$$\mathbf{x} = (x_1, x_2, x_3) \quad \text{ή} \quad \mathbf{u} = (x, y, z).$$

Το άθροισμα δύο διανυσμάτων στον $\mathbb{R}^n$ ορίζεται ως εξής:

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n).$$

Το γινόμενο αριθμού και διανύσματος στον $\mathbb{R}^n$ ορίζεται ως εξής:

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n).$$

Το αποτέλεσμα αυτών των πράξεων είναι διάνυσμα στον $\mathbb{R}^n$. Το μηδενικό διάνυσμα στον $\mathbb{R}^n$ είναι το

$$\mathbf{0} = (0, \dots, 0).$$

Αποδείξαμε ότι ο γεωμετρικός τρόπος πρόσθεσης δύο διανυσμάτων στον $\mathbb{R}^2$ είναι ο εξής: το $\mathbf{u} + \mathbf{v}$ είναι η διαγώνιος του παραλληλογράμμου που σχηματίζεται από τα $\mathbf{u}, \mathbf{v}$ (και τα τρία διανύσματα έχουν αρχή το σημείο $\mathbf{0}$). Αποδείξαμε ότι ο γεωμετρικός τρόπος πολλαπλασιασμού αριθμού και διανύσματος στον $\mathbb{R}^2$ είναι ο εξής: το $\lambda \mathbf{u}$ είναι το διάνυσμα στην ίδια κατεύθυνση, αν $\lambda > 0$, με το $\mathbf{u}$ και στην αντίθετη κατεύθυνση, αν $\lambda < 0$, με το $\mathbf{u}$ και με μήκος ίσο με $|\lambda|$ φορές το μήκος του $\mathbf{u}$. Το ότι ισχύουν τα ίδια και στον $\mathbb{R}^3$ το αφήσαμε ως άσκηση για το σπίτι.

Είδαμε ότι η ευθεία στον $\mathbb{R}^2$ ή στον $\mathbb{R}^3$ που παράγεται από ένα διάνυσμα $\mathbf{u} \neq \mathbf{0}$ στον $\mathbb{R}^2$ ή στον $\mathbb{R}^3$ αποτελείται από όλα τα διανύσματα $\lambda \mathbf{u}$, όπου το λ διατρέχει το $\mathbb{R}$. Επίσης, είδαμε ότι το επίπεδο στον $\mathbb{R}^3$ που παράγεται από δύο διανύσματα $\mathbf{u}, \mathbf{v}$ στον $\mathbb{R}^3$, τα οποία δεν είναι συνευθειακά, αποτελείται από όλα τα διανύσματα $\lambda \mathbf{u} + \mu \mathbf{v}$, όπου τα λ, μ διατρέχουν το $\mathbb{R}$.

Το μήκος ενός διανύσματος $\mathbf{x} = (x_1, \dots, x_n)$ ορίζεται να είναι το

$$\|\mathbf{x}\| = \sqrt{x_1^2 + \dots + x_n^2}.$$

Αποδείξαμε ότι για διανύσματα στον $\mathbb{R}^2$ και στον $\mathbb{R}^3$ ο τύπος αυτός του μήκους συμφωνεί με τον τύπο του μήκους που μαθαίνουμε στην γεωμετρία.

Αποδείξαμε τις βασικές ιδιότητες του μήκους:

$$\|\mathbf{x}\| \geq 0$$

$$\|\mathbf{x}\| = 0 \quad \text{αν και μόνο αν} \quad \mathbf{x} = \mathbf{0}$$

$$\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|.$$

Είδαμε ότι, αν $\mathbf{x} \neq \mathbf{0}$, τότε το διάνυσμα $\frac{1}{\|\mathbf{x}\|} \mathbf{x}$ έχει την ίδια κατεύθυνση με το $\mathbf{x}$ και μήκος ίσο με 1 (δηλαδή είναι μοναδιαίο).

Το εσωτερικό γινόμενο δύο διανυσμάτων $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$ ορίζεται να είναι το

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \dots + x_n y_n.$$

Το εσωτερικό γινόμενο διανυσμάτων είναι πάντοτε αριθμός.

Αποδείξαμε τις βασικές ιδιότητες του εσωτερικού γινομένου:

$$\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2$$

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$$

$$(\lambda \mathbf{x} + \mu \mathbf{y}) \cdot \mathbf{z} = \lambda(\mathbf{x} \cdot \mathbf{z}) + \mu(\mathbf{y} \cdot \mathbf{z})$$

$$\mathbf{x} \cdot (\lambda \mathbf{y} + \mu \mathbf{z}) = \lambda(\mathbf{x} \cdot \mathbf{y}) + \mu(\mathbf{x} \cdot \mathbf{z})$$

Αποδείξαμε την χρήσιμη ταυτότητα

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 \pm 2\mathbf{x} \cdot \mathbf{y} + \|\mathbf{y}\|^2.$$

Αποδείξαμε την ανισότητα του Cauchy

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

και, βάσει αυτής, αποδείξαμε την τριγωνική ανισότητα

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

Είδαμε ότι η τριγωνική ανισότητα για διανύσματα στον $\mathbb{R}^2$ και στον $\mathbb{R}^3$ εκφράζει το γνωστό από την γεωμετρία: το άθροισμα των μηκών δύο πλευρών ενός τριγώνου δεν είναι μεγαλύτερο από το μήκος της τρίτης πλευράς του.

Αποδείξαμε την εξής γεωμετρική ταυτότητα για τρίγωνα:

$$a^2 = b^2 + c^2 - 2bc \cos \theta,$$

όπου a, b, c είναι τα μήκη των τριών πλευρών ενός τριγώνου και θ είναι η κυρτή (δηλαδή $0 \leq \theta \leq \pi$) γωνία του τριγώνου η οποία βρίσκεται απέναντι από την πλευρά με μήκος a . Βάσει αυτού είδαμε το γεωμετρικό νόημα του εσωτερικού γινομένου διανυσμάτων στον $\mathbb{R}^2$ και στον $\mathbb{R}^3$, δηλαδή ότι

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta,$$

όπου θ είναι η κυρτή γωνία ανάμεσα στα μη-μηδενικά διανύσματα $\mathbf{u}, \mathbf{v}$. Έτσι έχουμε και έναν τύπο υπολογισμού της γωνίας θ :

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

Έτσι, για μη-μηδενικά διανύσματα στον $\mathbb{R}^2$ και στον $\mathbb{R}^3$, βλέπουμε ότι $\mathbf{u} \cdot \mathbf{v}$ αν και μόνο αν $\theta = \pi/2$. Με αυτό σαν κίνητρο, λέμε γενικότερα για διανύσματα $\mathbf{x}, \mathbf{y}$ στον $\mathbb{R}^n$ ότι είναι ορθογώνια ή κάθετα, και αυτό το συμβολίζουμε $\mathbf{x} \perp \mathbf{y}$, αν $\mathbf{x} \cdot \mathbf{y} = 0$. Δηλαδή

$$\mathbf{x} \perp \mathbf{y} \quad \text{αν και μόνο αν} \quad \mathbf{x} \cdot \mathbf{y} = 0.$$

Η κανονική βάση του $\mathbb{R}^n$ αποτελείται από τα διανύσματα

$$\mathbf{e}_1 = (1, 0, \dots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0), \quad \dots, \quad \mathbf{e}_n = (0, \dots, 0, 1).$$

Είδαμε ότι κάθε διάνυσμα στον $\mathbb{R}^n$ γράφεται σαν γραμμικός συνδυασμός των $\mathbf{e}_1, \dots, \mathbf{e}_n$ ως εξής:

$$(x_1, \dots, x_n) = x_1 \mathbf{e}_1 + \dots + x_n \mathbf{e}_n.$$

Στον $\mathbb{R}^2$ χρησιμοποιούμε και τα σύμβολα

$$\mathbf{i} = \mathbf{e}_1 = (1, 0), \quad \mathbf{j} = \mathbf{e}_2 = (0, 1),$$

οπότε $(x, y) = x\mathbf{i} + y\mathbf{j}$. Στον $\mathbb{R}^3$ χρησιμοποιούμε και τα σύμβολα

$$\mathbf{i} = \mathbf{e}_1 = (1, 0, 0), \quad \mathbf{j} = \mathbf{e}_2 = (0, 1, 0), \quad \mathbf{k} = \mathbf{e}_3 = (0, 0, 1),$$

οπότε $(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

Είδαμε ότι

$$\mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Το εξωτερικό γινόμενο διανυσμάτων $\mathbf{u}_1 = (x_1, y_1, z_1)$, $\mathbf{u}_2 = (x_2, y_2, z_2)$ στον $\mathbb{R}^3$ ορίζεται να είναι το

$$\begin{aligned} \mathbf{u}_1 \times \mathbf{u}_2 &= (y_1 z_2 - z_1 y_2, z_1 x_2 - x_1 z_2, x_1 y_2 - y_1 x_2) \\ &= (y_1 z_2 - z_1 y_2)\mathbf{i} - (x_1 z_2 - z_1 x_2)\mathbf{j} + (x_1 y_2 - y_1 x_2)\mathbf{k}. \end{aligned}$$

Το εξωτερικό γινόμενο ορίζεται μόνο για διανύσματα στον $\mathbb{R}^3$ και το αποτέλεσμα είναι διάνυσμα στον $\mathbb{R}^3$. Ενώ το εσωτερικό γινόμενο διανυσμάτων είναι αριθμός.

Γράφοντας $y_1 z_2 - z_1 y_2 = \det \begin{bmatrix} y_1 & z_1 \\ y_2 & z_2 \end{bmatrix}$, $x_1 z_2 - z_1 x_2 = \det \begin{bmatrix} x_1 & z_1 \\ x_2 & z_2 \end{bmatrix}$ και $x_1 y_2 - y_1 x_2 = \det \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}$, είδαμε ότι

$$\mathbf{u}_1 \times \mathbf{u}_2 = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{bmatrix}.$$

Αυτή η ταυτότητα έχει μόνο μνημονική αξία και όχι πραγματικό περιεχόμενο, διότι η πρώτη σειρά του 3×3 πίνακα δεν αποτελείται από αριθμούς.

Είδαμε τις βασικές ιδιότητες του εξωτερικού γινομένου:

$$\mathbf{u} \times \mathbf{u} = \mathbf{0},$$

$$\mathbf{u}_1 \times \mathbf{u}_2 = -\mathbf{u}_2 \times \mathbf{u}_1,$$

$$\mathbf{u}_1 \times (\lambda \mathbf{u}_2 + \mu \mathbf{u}_3) = \lambda(\mathbf{u}_1 \times \mathbf{u}_2) + \mu(\mathbf{u}_1 \times \mathbf{u}_3),$$

$$(\lambda \mathbf{u}_1 + \mu \mathbf{u}_2) \times \mathbf{u}_3 = \lambda(\mathbf{u}_1 \times \mathbf{u}_3) + \mu(\mathbf{u}_2 \times \mathbf{u}_3).$$

Αποδείξαμε τις δύο πρώτες ιδιότητες και αφήσαμε τις δύο τελευταίες σαν άσκηση για το σπίτι.

Το μεικτό γινόμενο τριών διανυσμάτων $\mathbf{u}_1 = (x_1, y_1, z_1)$, $\mathbf{u}_2 = (x_2, y_2, z_2)$, $\mathbf{u}_3 = (x_3, y_3, z_3)$ στον $\mathbb{R}^3$ ορίζεται να είναι το $(\mathbf{u}_1 \times \mathbf{u}_2) \cdot \mathbf{u}_3$ και είναι αριθμός. Είδαμε ότι

$$(\mathbf{u}_1 \times \mathbf{u}_2) \cdot \mathbf{u}_3 = \det \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}.$$

Βάσει αυτού του τύπου αποδείξαμε ότι $(\mathbf{u}_1 \times \mathbf{u}_2) \cdot \mathbf{u}_1 = 0$ και $(\mathbf{u}_1 \times \mathbf{u}_2) \cdot \mathbf{u}_2 = 0$. Δηλαδή

$$(\mathbf{u}_1 \times \mathbf{u}_2) \perp \mathbf{u}_1, \quad (\mathbf{u}_1 \times \mathbf{u}_2) \perp \mathbf{u}_2.$$

Άρα το $\mathbf{u}_1 \times \mathbf{u}_2$ είναι διάνυσμα ορθογώνιο στο επίπεδο που παράγεται από τα $\mathbf{u}_1, \mathbf{u}_2$ αν αυτά δεν είναι συνευθειακά.

Αν $\mathbf{u}_1, \mathbf{u}_2$ είναι διανύσματα στον $\mathbb{R}^2$ ή στον $\mathbb{R}^3$, αποδείξαμε ότι το εμβαδό E του παραλληλογράμμου που ορίζεται από τα δύο αυτά διανύσματα είναι ίσο με

$$E = \|\mathbf{u}_1\| \|\mathbf{u}_2\| \sin \theta,$$

όπου θ είναι η κυρτή γωνία ανάμεσα στα διανύσματα, και

$$E = \sqrt{\|\mathbf{u}_1\|^2 \|\mathbf{u}_2\|^2 - (\mathbf{u}_1 \cdot \mathbf{u}_2)^2}.$$

Ειδικά, αν τα $\mathbf{u}_1 = (x_1, y_1), \mathbf{u}_2 = (x_2, y_2)$ είναι διανύσματα στον $\mathbb{R}^2$, τότε είδαμε ότι

$$E = \left| \det \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \right|.$$

Αποδείξαμε ότι το μήκος του $\mathbf{u}_1 \times \mathbf{u}_2$ ισούται με

$$\|\mathbf{u}_1 \times \mathbf{u}_2\| = \sqrt{\|\mathbf{u}_1\|^2 \|\mathbf{u}_2\|^2 - (\mathbf{u}_1 \cdot \mathbf{u}_2)^2}$$

και άρα, σύμφωνα με τους προηγούμενους τύπους,

$$\|\mathbf{u}_1 \times \mathbf{u}_2\| = E = \|\mathbf{u}_1\| \|\mathbf{u}_2\| \sin \theta,$$

όπου E είναι το εμβαδό του παραλληλογράμμου που ορίζεται από τα δύο διανύσματα και θ είναι η κυρτή γωνία τους.

Τέλος, αποδείξαμε ότι ο όγκος V του παραλληλεπιπέδου που ορίζεται από τρία διανύσματα $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ στον $\mathbb{R}^3$ είναι ίσος με

$$V = |(\mathbf{u}_1 \times \mathbf{u}_2) \cdot \mathbf{u}_3|.$$

2η εβδομάδα.

Συζητήσαμε την παραμετρική εξίσωση ευθείας στον $\mathbb{R}^2$:

$$\mathbf{u} = \mathbf{u}_0 + t\mathbf{w}, \quad t \in \mathbb{R}$$

ή, ισοδύναμα,

$$\begin{cases} x = x_0 + t\xi, \\ y = y_0 + t\eta, \end{cases} \quad t \in \mathbb{R},$$

όπου $\mathbf{u} = (x, y)$ είναι το τυχόν σημείο της ευθείας, $\mathbf{u}_0 = (x_0, y_0)$ είναι δοσμένο σημείο της ευθείας και $\mathbf{w} = (\xi, \eta)$ είναι δοσμένο μη-μηδενικό διάνυσμα παράλληλο στην ευθεία. Το t είναι η παράμετρος.

Μια παραλλαγή αυτής της κατάστασης έχουμε με δύο δοσμένα σημεία της ευθείας, το $\mathbf{u}_0 = (x_0, y_0)$ και το $\mathbf{u}_1 = (x_1, y_1)$. Τότε το διάνυσμα $\mathbf{w} = \mathbf{u}_1 - \mathbf{u}_0 = (x_1 - x_0, y_1 - y_0)$ είναι παράλληλο στην ευθεία και η παραμετρική εξίσωση της ευθείας γράφεται

$$\mathbf{u} = \mathbf{u}_0 + t(\mathbf{u}_1 - \mathbf{u}_0), \quad t \in \mathbb{R}$$

ή, ισοδύναμα,

$$\begin{cases} x = x_0 + t(x_1 - x_0), \\ y = y_0 + t(y_1 - y_0), \end{cases} \quad t \in \mathbb{R}.$$

Τα εντελώς ανάλογα ισχύουν για ευθεία στον $\mathbb{R}^3$. Η παραμετρική εξίσωση ευθείας στον $\mathbb{R}^3$ είναι η

$$\mathbf{u} = \mathbf{u}_0 + t\mathbf{w}, \quad t \in \mathbb{R}$$

ή, ισοδύναμα,

$$\begin{cases} x = x_0 + t\xi, \\ y = y_0 + t\eta, \\ z = z_0 + t\zeta, \end{cases} \quad t \in \mathbb{R},$$

όπου $\mathbf{u} = (x, y, z)$ είναι το τυχόν σημείο της ευθείας, $\mathbf{u}_0 = (x_0, y_0, z_0)$ είναι δοσμένο σημείο της ευθείας και $\mathbf{w} = (\xi, \eta, \zeta)$ είναι δοσμένο μη-μηδενικό διάνυσμα παράλληλο στην ευθεία.

Όταν έχουμε δύο δοσμένα σημεία της ευθείας, το $\mathbf{u}_0 = (x_0, y_0, z_0)$ και το $\mathbf{u}_1 = (x_1, y_1, z_1)$, τότε το διάνυσμα $\mathbf{w} = \mathbf{u}_1 - \mathbf{u}_0 = (x_1 - x_0, y_1 - y_0, z_1 - z_0)$ είναι παράλληλο στην ευθεία και η παραμετρική εξίσωση της ευθείας γράφεται

$$\mathbf{u} = \mathbf{u}_0 + t(\mathbf{u}_1 - \mathbf{u}_0), \quad t \in \mathbb{R}$$

ή, ισοδύναμα,

$$\begin{cases} x = x_0 + t(x_1 - x_0), \\ y = y_0 + t(y_1 - y_0), \\ z = z_0 + t(z_1 - z_0), \end{cases} \quad t \in \mathbb{R}.$$

Η παραμετρική εξίσωση επιπέδου στον $\mathbb{R}^3$ είναι η

$$\mathbf{u} = \mathbf{u}_0 + t\mathbf{w}_1 + s\mathbf{w}_2, \quad t, s \in \mathbb{R}$$

ή, ισοδύναμα,

$$\begin{cases} x = x_0 + t\xi_1 + s\xi_2, \\ y = y_0 + t\eta_1 + s\eta_2, \\ z = z_0 + t\zeta_1 + s\zeta_2 \end{cases} \quad t, s \in \mathbb{R},$$

όπου $\mathbf{u} = (x, y, z)$ είναι το τυχόν σημείο του επιπέδου, $\mathbf{u}_0 = (x_0, y_0, z_0)$ είναι δοσμένο σημείο του επιπέδου και $\mathbf{w}_1 = (\xi_1, \eta_1, \zeta_1)$, $\mathbf{w}_2 = (\xi_2, \eta_2, \zeta_2)$ είναι δοσμένα μη-συνευθειακά διανύσματα παράλληλα στο επίπεδο. Τα t, s είναι οι παράμετροι.

Παραλλαγή αυτής της κατάστασης έχουμε με τρία δοσμένα μη-συνευθειακά σημεία του επιπέδου, το $\mathbf{u}_0 = (x_0, y_0, z_0)$ το $\mathbf{u}_1 = (x_1, y_1, z_1)$ και το $\mathbf{u}_2 = (x_2, y_2, z_2)$. Τότε τα διανύσματα $\mathbf{w}_1 = \mathbf{u}_1 - \mathbf{u}_0 = (x_1 - x_0, y_1 - y_0, z_1 - z_0)$, $\mathbf{w}_2 = (x_2 - x_0, y_2 - y_0, z_2 - z_0)$ είναι μη-συνευθειακά και παράλληλα στο επίπεδο και η παραμετρική εξίσωση του επιπέδου γράφεται

$$\mathbf{u} = \mathbf{u}_0 + t(\mathbf{u}_1 - \mathbf{u}_0) + s(\mathbf{u}_2 - \mathbf{u}_0), \quad t, s \in \mathbb{R}$$

ή, ισοδύναμα,

$$\begin{cases} x = x_0 + t(x_1 - x_0) + s(x_2 - x_0), \\ y = y_0 + t(y_1 - y_0) + s(y_2 - y_0), \\ z = z_0 + t(z_1 - z_0) + s(z_2 - z_0) \end{cases} \quad t, s \in \mathbb{R}.$$

Κατόπιν, μελετήσαμε καρτεσιανές εξισώσεις ευθειών και επιπέδων.

Η **καρτεσιανή εξίσωση** ευθείας στον $\mathbb{R}^2$ γράφεται

$$(\mathbf{u} - \mathbf{u}_0) \cdot \mathbf{v} = 0,$$

όπου $\mathbf{u} = (x, y)$ είναι το τυχόν σημείο της ευθείας, $\mathbf{u}_0 = (x_0, y_0)$ είναι δοσμένο σημείο της ευθείας και $\mathbf{v} = (a, b)$ είναι δοσμένο μη-μηδενικό διάνυσμα κάθετο στην ευθεία. Ισοδύναμα:

$$ax + by = ax_0 + by_0.$$

Οπότε, αν θεωρήσουμε την σταθερά $c = ax_0 + by_0$, η καρτεσιανή εξίσωση γράφεται

$$ax + by = c.$$

Μία ευθεία στον $\mathbb{R}^3$ περιγράφεται με δύο καρτεσιανές εξισώσεις:

$$\begin{cases} (\mathbf{u} - \mathbf{u}_0) \cdot \mathbf{v}_1 = 0 \\ (\mathbf{u} - \mathbf{u}_0) \cdot \mathbf{v}_2 = 0 \end{cases}$$

όπου $\mathbf{u} = (x, y, z)$ είναι το τυχόν σημείο της ευθείας, $\mathbf{u}_0 = (x_0, y_0, z_0)$ είναι δοσμένο σημείο της ευθείας και $\mathbf{v}_1 = (a_1, b_1, c_1)$, $\mathbf{v}_2 = (a_2, b_2, c_2)$ είναι δοσμένα μη-συνευθειακά διανύσματα κάθετα στην ευθεία (οπότε η ευθεία είναι κάθετη στο επίπεδο που παράγεται από τα $\mathbf{v}_1, \mathbf{v}_2$). Ισοδύναμα:

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$$

όπου $d_1 = a_1x_0 + b_1y_0 + c_1z_0$, $d_2 = a_2x_0 + b_2y_0 + c_2z_0$.

Ένα επίπεδο στον $\mathbb{R}^3$ περιγράφεται με μία καρτεσιανή εξίσωση:

$$(\mathbf{u} - \mathbf{u}_0) \cdot \mathbf{v} = 0,$$

όπου $\mathbf{u} = (x, y, z)$ είναι το τυχόν σημείο του επιπέδου, $\mathbf{u}_0 = (x_0, y_0, z_0)$ είναι δοσμένο σημείο του επιπέδου και $\mathbf{v} = (a, b, c)$ είναι δοσμένο μη-μηδενικό διάνυσμα κάθετο στο επίπεδο. Ισοδύναμα:

$$ax + by + cz = d,$$

όπου $d = ax_0 + by_0 + cz_0$.

Κάναμε μερικές παρατηρήσεις:

A. Για να περιγράψουμε μια ευθεία *παραμετρικά*, είτε στον $\mathbb{R}^2$ είτε στον $\mathbb{R}^3$, χρειαζόμαστε μία παράμετρο. Αυτό έχει να κάνει με το ότι η ευθεία είναι μονοδιάστατο “αντικείμενο”. Για να περιγράψουμε ένα επίπεδο *παραμετρικά* στον $\mathbb{R}^3$ χρειαζόμαστε δύο παραμέτρους. Αυτό έχει να κάνει με το ότι το επίπεδο είναι διδιάστατο “αντικείμενο”.

B. Παραμετρικές αναπαραστάσεις σχετίζονται με παραλληλία. Πράγματι, η παραμετρική αναπαράσταση ευθείας, είτε στον $\mathbb{R}^2$ είτε στον $\mathbb{R}^3$, προκύπτει από ένα διάνυσμα παράλληλο στην ευθεία και η παραμετρική αναπαράσταση επιπέδου στον $\mathbb{R}^3$ προκύπτει από δύο διανύσματα παράλληλα στο επίπεδο (δηλαδή από ένα επίπεδο παράλληλο στο επίπεδο).

Γ. Καρτεσιανές αναπαραστάσεις σχετίζονται με καθετότητα. Πράγματι, η καρτεσιανή αναπαράσταση ευθείας στον $\mathbb{R}^2$ προκύπτει από ένα διάνυσμα (δηλαδή, ευθεία) κάθετο στην ευθεία. Η καρτεσιανή αναπαράσταση ευθείας στον $\mathbb{R}^3$ προκύπτει από δύο διανύσματα (δηλαδή, επίπεδο) κάθετα στην ευθεία. Η καρτεσιανή αναπαράσταση επιπέδου στον $\mathbb{R}^3$ προκύπτει από ένα διάνυσμα (δηλαδή, ευθεία) κάθετο στην ευθεία.

Δ. Μία ευθεία στον $\mathbb{R}^2$ περιγράφεται από μία καρτεσιανή εξίσωση (από το κάθετο διάνυσμα ή, ισοδύναμα, από την κάθετη ευθεία). Μία ευθεία στον $\mathbb{R}^3$ περιγράφεται από δύο καρτεσιανές εξισώσεις (από τα δύο κάθετα διανύσματα ή, ισοδύναμα, από το κάθετο επίπεδο). Ένα επίπεδο στον $\mathbb{R}^3$ περιγράφεται από μία καρτεσιανή εξίσωση (από το κάθετο διάνυσμα ή, ισοδύναμα, από την κάθετη ευθεία).

Ε. Όπως είπαμε στο A, ο αριθμός k των παραμέτρων που χρειάζονται για την περιγραφή ενός “αντικειμένου” (ευθείας ή επιπέδου) είναι ίδιο με την διάσταση αυτού του “αντικειμένου”. Ο αριθμός m των καρτεσιανών εξισώσεων που χρειάζονται για την περιγραφή ενός “αντικειμένου” είναι ίδιο με την διάσταση του κάθετου “αντικειμένου”. Επειδή η διάσταση του “αντικειμένου” και η διάσταση του κάθετου “αντικειμένου” έχουν άθροισμα ίσο με τη συνολική διάσταση n του περιβάλλοντος χώρου, έχουμε ότι

$$k + m = n.$$

(Στις περιπτώσεις που εξετάσαμε, έχουμε: $1 + 1 = 2$, $1 + 2 = 3$, $2 + 1 = 3$.)

Η **προβολή** διανύσματος $\mathbf{u}$ στην ευθεία που παράγεται από ένα μη-μηδενικό διάνυσμα $\mathbf{v}$ είναι κάποιο πολλαπλάσιο $t\mathbf{v}$ του $\mathbf{v}$ τέτοιο ώστε η διαφορά $\mathbf{u} - t\mathbf{v}$ να είναι κάθετη στην ευθεία ή, ισοδύναμα, κάθετο στο $\mathbf{v}$ το οποίο παράγει την ευθεία. Από την εξίσωση

$$(\mathbf{u} - t\mathbf{v}) \cdot \mathbf{v} = 0$$

προκύπτει

$$t = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2}$$

και άρα η προβολή είναι το διάνυσμα

$$\text{Pr}_{\mathbf{v}}(\mathbf{u}) = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{v}\|^2} \mathbf{v}.$$

Η **προβολή** διανύσματος $\mathbf{u}$ στο επίπεδο που παράγεται από δύο μη-συνευθειακά διανύσματα $\mathbf{v}_1, \mathbf{v}_2$ είναι κάποιος γραμμικός συνδυασμός $t_1\mathbf{v}_1 + t_2\mathbf{v}_2$ τέτοιος ώστε η διαφορά $\mathbf{u} - (t_1\mathbf{v}_1 + t_2\mathbf{v}_2)$ να είναι κάθετη στο επίπεδο ή, ισοδύναμα, κάθετη στα $\mathbf{v}_1, \mathbf{v}_2$ τα οποία παράγουν το επίπεδο. Από το σύστημα των εξισώσεων

$$\begin{cases} (\mathbf{u} - t_1\mathbf{v}_1 - t_2\mathbf{v}_2) \cdot \mathbf{v}_1 = 0 \\ (\mathbf{u} - t_1\mathbf{v}_1 - t_2\mathbf{v}_2) \cdot \mathbf{v}_2 = 0 \end{cases}$$

προκύπτει

$$t_1 = \frac{\det \begin{bmatrix} \mathbf{u} \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_1 \\ \mathbf{u} \cdot \mathbf{v}_2 & \mathbf{v}_2 \cdot \mathbf{v}_2 \end{bmatrix}}{\det \begin{bmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_1 \\ \mathbf{v}_1 \cdot \mathbf{v}_2 & \mathbf{v}_2 \cdot \mathbf{v}_2 \end{bmatrix}}, \quad t_2 = \frac{\det \begin{bmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{u} \cdot \mathbf{v}_1 \\ \mathbf{v}_1 \cdot \mathbf{v}_2 & \mathbf{u} \cdot \mathbf{v}_2 \end{bmatrix}}{\det \begin{bmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_1 \\ \mathbf{v}_1 \cdot \mathbf{v}_2 & \mathbf{v}_2 \cdot \mathbf{v}_2 \end{bmatrix}}.$$

Άρα η προβολή είναι το διάνυσμα

$$\text{Pr}_{\mathbf{v}_1, \mathbf{v}_2}(\mathbf{u}) = t_1 \mathbf{v}_1 + t_2 \mathbf{v}_2,$$

όπου τα t_1, t_2 δίνονται από τους προηγούμενους τύπους.

Κατόπιν αποδείξαμε ότι στον $\mathbb{R}^2$ η απόσταση D σημείου $\mathbf{u}_0 = (x_0, y_0)$ από ευθεία με καρτεσιανή εξίσωση $ax + by = c$ δίνεται από τον τύπο

$$D = \frac{|ax_0 + by_0 - c|}{\sqrt{a^2 + b^2}}.$$

Ο αντίστοιχος τύπος στον $\mathbb{R}^3$ για την απόσταση D σημείου $\mathbf{u}_0 = (x_0, y_0, z_0)$ από επίπεδο με καρτεσιανή εξίσωση $ax + by + cz = d$ δίνεται από τον τύπο

$$D = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Να θυμηθούμε μερικά χρήσιμα στοιχεία από τη Γραμμική Άλγεβρα.

Θεωρούμε m διανύσματα στον χώρο $\mathbb{R}^n$:

$$\begin{aligned} \mathbf{x}_1 &= (x_{11}, \dots, x_{1j}, \dots, x_{1n}) \\ &\vdots \\ \mathbf{x}_i &= (x_{i1}, \dots, x_{ij}, \dots, x_{in}) \\ &\vdots \\ \mathbf{x}_m &= (x_{m1}, \dots, x_{mj}, \dots, x_{mn}) \end{aligned}$$

Τα διανύσματα αυτά παράγουν έναν γραμμικό υπόχωρο, έστω V , του $\mathbb{R}^n$. Εξ ορισμού, ο γραμμικός υπόχωρος V που παράγεται από τα $\mathbf{x}_1, \dots, \mathbf{x}_m$ αποτελείται από όλους τους γραμμικούς συνδυασμούς αυτών των m διανυσμάτων. Δηλαδή

$$V = \{\lambda_1 \mathbf{x}_1 + \dots + \lambda_m \mathbf{x}_m \mid \lambda_1, \dots, \lambda_m \in \mathbb{R}\}.$$

Ερώτημα 1: Ποιά είναι η διάσταση του V ;

Ο ορισμός της διάστασης του V είναι ο εξής: το πλήθος των στοιχείων μιας βάσης του V . Δηλαδή, αν βρούμε κάποια στοιχεία $\mathbf{y}_1, \dots, \mathbf{y}_k$ του V τα οποία παράγουν τον V και ταυτόχρονα είναι και γραμμικώς ανεξάρτητα, τότε αυτά αποτελούν βάση του V και επομένως η διάσταση του V είναι ίση με k .

Επεξήγηση: Το ότι τα $\mathbf{y}_1, \dots, \mathbf{y}_k$ παράγουν τον V σημαίνει ότι ο V αποτελείται από όλους τους γραμμικούς συνδυασμούς των στοιχείων αυτών. Το ότι τα $\mathbf{y}_1, \dots, \mathbf{y}_k$ είναι γραμμικώς ανεξάρτητα σημαίνει ότι κανένα από αυτά δεν είναι γραμμικός συνδυασμός των υπολοίπων. Τα $\mathbf{x}_1, \dots, \mathbf{x}_m$ είναι (γιατί;) στοιχεία του V και παράγουν τον V . Αν είναι και γραμμικώς ανεξάρτητα, τότε αποτελούν βάση του V και άρα η διάσταση του V είναι ίση με m . Αν δεν είναι γραμμικώς ανεξάρτητα, τότε δεν αποτελούν βάση του V . Σ' αυτήν την περίπτωση ο V θα έχει μια άλλη βάση με λιγότερα στοιχεία.

Ερώτημα 2: Πώς βρίσκουμε μια βάση του V και άρα και την διάστασή του;

Υπάρχουν δύο τρόποι (ουσιαστικά ισοδύναμοι).

Πρώτος τρόπος εύρεσης βάσης του V . Αν τα $\mathbf{x}_1, \dots, \mathbf{x}_m$ (το οποία, όπως είπαμε, παράγουν τον V) είναι γραμμικώς ανεξάρτητα τότε αποτελούν βάση του V και η διάσταση του V είναι ίση με m . Αν τα $\mathbf{x}_1, \dots, \mathbf{x}_m$ δεν είναι γραμμικώς ανεξάρτητα, τότε κάποιο από αυτά είναι γραμμικός συνδυασμός των υπολοίπων. Ας υποθέσουμε ότι το $\mathbf{x}_m$ είναι γραμμικός συνδυασμός των $\mathbf{x}_1, \dots, \mathbf{x}_{m-1}$. (Αν δεν είναι το $\mathbf{x}_m$, μπορούμε να αριθμήσουμε διαφορετικά τα διανύσματα ώστε το τελευταίο να είναι γραμμικός συνδυασμός των προηγούμενων.) Τότε μπορούμε να παραλείψουμε το $\mathbf{x}_m$ και τα υπόλοιπα, $\mathbf{x}_1, \dots, \mathbf{x}_{m-1}$, παράγουν, επίσης, τον V . Αν αυτά είναι γραμμικώς ανεξάρτητα, τότε αποτελούν βάση του V και άρα η διάσταση του V είναι ίση με $m - 1$. Αν τα $\mathbf{x}_1, \dots, \mathbf{x}_{m-1}$ δεν είναι γραμμικώς ανεξάρτητα, τότε κάποιο από αυτά είναι γραμμικός συνδυασμός των υπολοίπων. Ας υποθέσουμε ότι το $\mathbf{x}_{m-1}$ είναι γραμμικός συνδυασμός των $\mathbf{x}_1, \dots, \mathbf{x}_{m-2}$. (Αν δεν είναι το $\mathbf{x}_{m-1}$, μπορούμε να αριθμήσουμε διαφορετικά τα διανύσματα ώστε το τελευταίο να είναι γραμμικός συνδυασμός των προηγούμενων.) Τότε μπορούμε να παραλείψουμε το $\mathbf{x}_{m-1}$ και τα υπόλοιπα, $\mathbf{x}_1, \dots, \mathbf{x}_{m-2}$, παράγουν, επίσης, τον V . Αν αυτά είναι γραμμικώς ανεξάρτητα, τότε αποτελούν βάση του V και άρα η διάσταση του V είναι ίση με $m - 2$. Συνεχίζουμε αυτήν την διαδικασία, παραλείποντας διαδοχικά τα $\mathbf{x}_m, \mathbf{x}_{m-1}, \dots$ (και αλλάζοντας αρίθμηση αν χρειάζεται) μέχρι να βρούμε κάποια στιγμή $\mathbf{x}_1, \dots, \mathbf{x}_k$ τα οποία παράγουν τον V και είναι γραμμικώς ανεξάρτητα. Αυτά αποτελούν βάση του V και η διάστασή του θα είναι ίση με k .

Δεύτερος (και προτιμότερος) τρόπος εύρεσης βάσης του V . Εφαρμόζουμε τη διαδικασία απαλοιφής του Gauss στον $m \times n$ πίνακα A που σχηματίζεται με γραμμές τις n -άδες $\mathbf{x}_1, \dots, \mathbf{x}_m$:

$$A = \begin{bmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1n} \\ \vdots & & \vdots & & \vdots \\ x_{i1} & \dots & x_{ij} & \dots & x_{in} \\ \vdots & & \vdots & & \vdots \\ x_{m1} & \dots & x_{mj} & \dots & x_{mn} \end{bmatrix}$$

Κάνοντας τις γνωστές πράξεις με τις γραμμές του πίνακα A καταλήγουμε σε έναν $m \times n$ πίνακα B ο οποίος έχει τη μορφή

$$B = \begin{bmatrix} 0 & \dots & 0 & a_1 & * & \dots & * & * & * & \dots & * & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & a_2 & * & \dots & * & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & a_k & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}$$

Η πρώτη γραμμή έχει κάποια αρχικά 0 (μπορεί κανένα), το πρώτο μη-μηδενικό στοιχείο a_1 και ακολουθούν τυχόντες αριθμοί. Η δεύτερη γραμμή έχει κάποια αρχικά 0, το πρώτο μη-μηδενικό a_2 το οποίο είναι δεξιά του a_1 και ακολουθούν τυχόντες αριθμοί. Συνεχίζουμε μέχρι και την k -οστή γραμμή, η οποία έχει κάποια αρχικά 0, το πρώτο μη-μηδενικό a_k το οποίο είναι δεξιά του a_{k-1} της αμέσως προηγούμενης γραμμής και ακολουθούν τυχόντες αριθμοί. Οι υπόλοιπες γραμμές (μπορεί καμία) αποτελούνται μόνο από 0. Οι αριθμοί $a_1, a_2, \dots, a_k$ είναι οι λεγόμενοι οδηγοί των αντίστοιχων γραμμών. Όταν καταλήξουμε σε έναν τέτοιο πίνακα B , τότε οι πρώτες k γραμμές είναι διανύσματα στον $\mathbb{R}^n$ τα οποία παράγουν τον V και είναι γραμμικώς ανεξάρτητα και άρα αποτελούν βάση του V . Επομένως έχουμε βρει μία βάση του V και τη διάστασή του η οποία είναι ίση με k .

Παρατήρηση: Αν $m = 2$, δηλαδή αν έχουμε δύο διανύσματα, τα $\mathbf{x}_1, \mathbf{x}_2$, τότε δεν χρειάζεται να κάνουμε όλη αυτήν τη διαδικασία. Είναι εύκολο να δούμε “με το μάτι” αν ένα από τα δύο διανύσματα είναι πολλαπλάσιο του άλλου. Αν κανένα από τα δύο διανύσματα δεν είναι πολλαπλάσιο του άλλου (δηλαδή δεν είναι “συνευθειακά”), τότε αυτά είναι γραμμικώς ανεξάρτητα και άρα αποτελούν

βάση του χώρου V τον οποίο παράγουν. Δηλαδή, τα δύο διανύσματα παράγουν χώρο διάστασης 2. (Ειδική περίπτωση: Ο V είναι ολόκληρος ο $\mathbb{R}^2$ αν τα $\mathbf{x}_1, \mathbf{x}_2$ είναι στον $\mathbb{R}^2$ ή ο V είναι ένα επίπεδο στον $\mathbb{R}^3$ αν τα $\mathbf{x}_1, \mathbf{x}_2$ είναι στον $\mathbb{R}^3$.) Αν ένα από τα δύο διανύσματα είναι πολλαπλάσιο του άλλου, π.χ. $\mathbf{x}_2 = \lambda \mathbf{x}_1$, τότε το $\mathbf{x}_1$ αποτελεί βάση του χώρου V που παράγουν τα δύο διανύσματα (δηλαδή, από μόνο του παράγει τον V). Άρα ο V έχει διάσταση 1. (Ειδική περίπτωση: Ο V είναι μία ευθεία στον $\mathbb{R}^2$ αν τα $\mathbf{x}_1, \mathbf{x}_2$ είναι στον $\mathbb{R}^2$ ή ο V είναι μία ευθεία στον $\mathbb{R}^3$ αν τα $\mathbf{x}_1, \mathbf{x}_2$ είναι στον $\mathbb{R}^3$.) Η διαδικασία με την μέθοδο του Gauss έχει αξία όταν έχουμε $m \geq 3$ διανύσματα, οπότε δεν μπορούμε να δούμε εύκολα “με το μάτι” αν ένα από τα $\mathbf{x}_1, \dots, \mathbf{x}_m$ είναι γραμμικός συνδυασμός των υπολοίπων.

Μελετήσαμε τις **πολικές συντεταγμένες** στον $\mathbb{R}^2$.

Σε ένα σημείο στον $\mathbb{R}^2$ με ζεύγος καρτεσιανών συντεταγμένων (x, y) αντιστοιχίζουμε το ζεύγος πολικών συντεταγμένων (r, θ) . Το r είναι η απόσταση αυτού του σημείου από την αρχή των αξόνων και το θ είναι η γωνία ανάμεσα στο διάνυσμα (x, y) και στον θετικό x -ημιάξονα. Για την απόσταση r ισχύει προφανώς $r \geq 0$ και για την γωνία θ ισχύει $0 \leq \theta \leq 2\pi$. Η γωνία θ μετριέται ξεκινώντας από τον θετικό x -ημιάξονα και πηγαίνοντας στο διάνυσμα (x, y) με την **θετική φορά περιστροφής** γύρω από την αρχή των αξόνων, δηλαδή την αντίθετη προς την φορά περιστροφής των δεικτών του ρολογιού. (Προσοχή: η γωνία θ δεν είναι πάντοτε η κυρτή γωνία ανάμεσα στο διάνυσμα (x, y) και στον θετικό x -ημιάξονα.)

Είδαμε ότι οι τύποι αλλαγής καρτεσιανών-πολικών συντεταγμένων είναι οι

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

Επίσης, έχουμε και τους τύπους $r = \sqrt{x^2 + y^2}$ και $\theta = \arctan \frac{y}{x}$. Ο δεύτερος τύπος ισχύει για $x, y > 0$ (ή ισοδύναμα $0 < \theta < \frac{\pi}{2}$).

Πρώτη παρατήρηση: Στα σημεία του θετικού x -ημιάξονα αντιστοιχίζονται δύο τιμές του θ , το 0 και το 2π .

Δεύτερη παρατήρηση: Στο σημείο $(0, 0)$ αντιστοιχίζεται το $r = 0$. Αλλά κάποιοι συγγραφείς στο σημείο αυτό δεν αντιστοιχίζουν καμία τιμή του θ ενώ άλλοι συγγραφείς αντιστοιχίζουν όλες τις τιμές του θ .

Κατόπιν μελετήσαμε τις **κυλινδρικές συντεταγμένες** στον $\mathbb{R}^3$.

Σε ένα σημείο στον $\mathbb{R}^3$ με τριάδα καρτεσιανών συντεταγμένων (x, y, z) αντιστοιχίζουμε τριάδα κυλινδρικών συντεταγμένων (r, θ, z) . Το (r, θ) είναι το ζεύγος πολικών συντεταγμένων που αντιστοιχίζεται στο σημείο $(x, y, 0)$ στο xy -επίπεδο. Το σημείο αυτό είναι η κατακόρυφη προβολή του (x, y, z) στο xy -επίπεδο. Άρα έχουμε $r \geq 0$ και $0 \leq \theta \leq 2\pi$. Για το z δεν υπάρχει περιορισμός. Είδαμε, λοιπόν, ότι οι τύποι αλλαγής καρτεσιανών-κυλινδρικών συντεταγμένων είναι οι

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \\ z = z \end{cases}$$

Ισχύει πάλι ο τύπος $r = \sqrt{x^2 + y^2}$ καθώς και ο $\theta = \arctan \frac{y}{x}$ για $x, y > 0$ (ή ισοδύναμα $0 < \theta < \frac{\pi}{2}$).

Πρώτη παρατήρηση: Στα σημεία του δεξιού xz -ημιεπιπέδου αντιστοιχίζονται δύο τιμές του θ , το 0 και το 2π .

Δεύτερη παρατήρηση: Στα σημεία του z -άξονα αντιστοιχίζεται το $r = 0$. Τώρα, κάποιοι συγγραφείς στα σημεία αυτά δεν αντιστοιχίζουν καμία τιμή του θ ενώ άλλοι συγγραφείς αντιστοιχίζουν όλες τις τιμές του θ .

Τέλος μελετήσαμε τις **σφαιρικές συντεταγμένες** στον $\mathbb{R}^3$.

Σε ένα σημείο στον $\mathbb{R}^3$ με τριάδα καρτεσιανών συντεταγμένων (x, y, z) αντιστοιχίζουμε τριάδα

σφαιρικών συντεταγμένων (ρ, θ, ϕ) . Το ρ είναι η απόσταση του σημείου (x, y, z) από την αρχή των αξόνων, το θ είναι η δεύτερη πολική συντεταγμένη του σημείου $(x, y, 0)$ στο xy -επίπεδο και το ϕ είναι η κυρτή γωνία ανάμεσα στο διάνυσμα (x, y, z) και στον θετικό z -ημιάξονα. Άρα έχουμε $\rho \geq 0$ και $0 \leq \theta \leq 2\pi$ και $0 \leq \phi \leq \pi$.

Είδαμε ότι οι τύποι αλλαγής καρτεσιανών-σφαιρικών συντεταγμένων είναι οι

$$\begin{cases} x = \rho \cos \theta \sin \phi \\ y = \rho \sin \theta \sin \phi \\ z = \rho \cos \phi \end{cases}$$

Έχουμε και τον τύπο $\rho = \sqrt{x^2 + y^2 + z^2}$ καθώς και τον $\theta = \arctan \frac{y}{x}$ για $x, y > 0$. Επίσης, τον τύπο $\phi = \arccos \frac{z}{\rho}$.

Πρώτη παρατήρηση: Στα σημεία του δεξιού xz -ημιεπιπέδου αντιστοιχίζονται δύο τιμές του θ , το 0 και το 2π .

Δεύτερη παρατήρηση: Στα σημεία του z -άξονα κάποιοι συγγραφείς δεν αντιστοιχίζουν καμία τιμή του θ ενώ άλλοι συγγραφείς αντιστοιχίζουν όλες τις τιμές του θ . Επίσης, στο σημείο $(0, 0, 0)$ κάποιοι συγγραφείς δεν αντιστοιχίζουν καμία τιμή του ϕ ενώ άλλοι συγγραφείς αντιστοιχίζουν όλες τις τιμές του ϕ .

3η εβδομάδα.

Συζητήσαμε για μερικά χαρακτηριστικά σχήματα στον $\mathbb{R}^2$ και στον $\mathbb{R}^3$.

Σχήματα στο επίπεδο.

1. Η καρτεσιανή εξίσωση

$$x^2 + y^2 = r_0^2,$$

όπου το $r_0 > 0$ είναι δοσμένος αριθμός, ορίζει τον **κύκλο** κέντρου $(0, 0)$ και ακτίνας r_0 . Η πολική εξίσωση του ίδιου κύκλου είναι η

$$r = r_0$$

(το θ είναι ελεύθερο στο διάστημα $[0, 2\pi]$).

2. Η καρτεσιανή εξίσωση

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1,$$

όπου τα $a, b > 0$ είναι δοσμένοι αριθμοί, ορίζει την **έλλειψη** με κέντρο $(0, 0)$ και ημιάξονες a, b . Αν $a = b = r_0$ παίρνουμε ως ειδική περίπτωση τον κύκλο ακτίνας r_0 .

3. Η καρτεσιανή εξίσωση

$$y = x^2$$

ορίζει την γνωστή **παραβολή** η οποία είναι συμμετρική ως προς τον y -άξονα, έχει κορυφή το $(0, 0)$ και ανοίγει προς τα πάνω. Με εναλλαγή των x, y προκύπτει η εξίσωση $x = y^2$ η οποία ορίζει την παραβολή, η οποία είναι συμμετρική ως προς τον x -άξονα, έχει κορυφή το $(0, 0)$ και ανοίγει προς τα δεξιά. Είδαμε πώς από αυτές τις δύο παραβολές προκύπτουν οι γενικότερες παραβολές με καρτεσιανές εξισώσεις

$$y = ax^2 + bx + c, \quad x = ay^2 + by + c,$$

με $a \neq 0$.

4. Η καρτεσιανή εξίσωση

$$x^2 - y^2 = c,$$

όπου το c είναι δοσμένος αριθμός, ορίζει **υπερβολή**.

Διακρίναμε περιπτώσεις.

Αν $c = 0$, το σχήμα που προκύπτει είναι η ένωση των δύο κύριων διαγωνίων του xy -επιπέδου.

Αν $c > 0$, προκύπτουν δύο κλάδοι υπερβολής. Και οι δύο κλάδοι είναι συμμετρικοί ως προς τον x -άξονα. Ο ένας έχει κορυφή το $(\sqrt{c}, 0)$ και ανοίγει προς τα δεξιά και ο άλλος έχει κορυφή το $(-\sqrt{c}, 0)$ και ανοίγει προς τα αριστερά. Και οι δύο κλάδοι έχουν ασύμπτωτες τις κύριες διαγώνιες. Οι δύο κλάδοι είναι μεταξύ τους συμμετρικοί ως προς τον y -άξονα.

Αν $c < 0$, προκύπτουν πάλι δύο κλάδοι υπερβολής. Και οι δύο κλάδοι είναι συμμετρικοί ως προς τον y -άξονα. Ο ένας έχει κορυφή το $(0, \sqrt{c})$ και ανοίγει προς τα πάνω και ο άλλος έχει κορυφή το $(0, -\sqrt{c})$ και ανοίγει προς τα κάτω. Και οι δύο κλάδοι έχουν ασύμπτωτες τις κύριες διαγώνιες. Οι δύο κλάδοι είναι μεταξύ τους συμμετρικοί ως προς τον x -άξονα.

Είδαμε πώς μεταβάλλονται αυτές οι υπερβολές όταν μεταβάλλεται το c .

5. Η καρτεσιανή εξίσωση

$$xy = c,$$

όπου το c είναι δοσμένος αριθμός, ορίζει **υπερβολή**.

Διακρίναμε περιπτώσεις.

Αν $c = 0$, το σχήμα που προκύπτει είναι η ένωση των δύο αξόνων: του x -άξονα και του y -άξονα.

Αν $c > 0$, προκύπτουν δύο κλάδοι υπερβολής. Και οι δύο κλάδοι είναι συμμετρικοί ως προς την κύρια διαγώνιο $y = x$. Ο ένας έχει κορυφή το $(\sqrt{c}, \sqrt{c})$ και ανοίγει προς τα δεξιά και πάνω και ο άλλος έχει κορυφή το $(-\sqrt{c}, -\sqrt{c})$ και ανοίγει προς τα αριστερά και κάτω. Και οι δύο κλάδοι έχουν ασύμπτωτες τους δύο άξονες. Οι δύο κλάδοι είναι μεταξύ τους συμμετρικοί ως προς την κύρια διαγώνιο $y = -x$.

Αν $c < 0$, προκύπτουν πάλι δύο κλάδοι υπερβολής. Και οι δύο κλάδοι είναι συμμετρικοί ως προς την κύρια διαγώνιο $y = -x$. Ο ένας έχει κορυφή το $(-\sqrt{c}, \sqrt{c})$ και ανοίγει προς τα αριστερά και πάνω και ο άλλος έχει κορυφή το $(\sqrt{c}, -\sqrt{c})$ και ανοίγει προς τα δεξιά και κάτω. Και οι δύο κλάδοι έχουν ασύμπτωτες τους δύο άξονες. Οι δύο κλάδοι είναι μεταξύ τους συμμετρικοί ως προς την κύρια διαγώνιο $y = x$.

Είδαμε πώς μεταβάλλονται αυτές οι υπερβολές όταν μεταβάλλεται το c .

Μελετήσαμε την σχέση ανάμεσα στις υπερβολές της περίπτωσης 4 και στις υπερβολές της περίπτωσης 5. Είδαμε ότι η καρτεσιανή εξίσωση $x^2 - y^2 = c$ γράφεται $(x - y)(x + y) = c$ και ότι, αν θέσουμε $\xi = x - y$ και $\eta = x + y$, τότε η εξίσωση γράφεται $\xi\eta = c$. Βάσει αυτού είδαμε ότι οι υπερβολές της περίπτωσης 4 προκύπτουν από τις υπερβολές της περίπτωσης 5 (και αντιστρόφως) με στροφή κατά γωνία $\frac{\pi}{4}$.

Σχήματα στον χώρο.

1. Η καρτεσιανή εξίσωση

$$x^2 + y^2 + z^2 = \rho_0^2,$$

όπου το $\rho_0 > 0$ είναι δοσμένος αριθμός, ορίζει την **σφαίρα** κέντρου $(0, 0, 0)$ και ακτίνας ρ_0 . Η σφαιρική εξίσωση της ίδιας σφαίρας είναι η

$$\rho = \rho_0$$

(τα θ, ϕ είναι ελεύθερα στα διαστήματα $[0, 2\pi]$, $[0, \pi]$ αντιστοίχως). Η κυλινδρική εξίσωση της ίδιας σφαίρας είναι η

$$r^2 + z^2 = \rho^2$$

(το θ είναι ελεύθερο στο διάστημα $[0, 2\pi]$).

2. Η καρτεσιανή εξίσωση

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1,$$

όπου τα $a, b, c > 0$ είναι δοσμένοι αριθμοί, ορίζει το **ελλειψοειδές** με κέντρο $(0, 0, 0)$ και ημιάξονες a, b, c . Αν $a = b = c = \rho_0$ παίρνουμε ως ειδική περίπτωση την σφαίρα ακτίνας ρ_0 .

Είδαμε με πολλή λεπτομέρεια ότι οι τομές του ελλειψοειδούς με επίπεδα παράλληλα στο yz -επίπεδο (δηλαδή κάθετα στον x -άξονα) είναι ελλείψεις. Πιο συγκεκριμένα, αν $-a < x_0 < a$, η τομή του ελλειψοειδούς με το επίπεδο με εξίσωση $x = x_0$ ικανοποιεί την εξίσωση

$$\left(\frac{y}{\kappa b}\right)^2 + \left(\frac{z}{\kappa c}\right)^2 = 1,$$

όπου $\kappa = \sqrt{1 - \left(\frac{x_0}{a}\right)^2} > 0$ (και το x είναι σταθερό: $x = x_0$). Όταν $x_0 = 0$, η αντίστοιχη έλλειψη είναι πάνω στο yz -επίπεδο. Όταν το x_0 αυξάνεται στο διάστημα $[0, a)$, η έλλειψη μεταφέρεται παράλληλα προς τα δεξιά και μικραίνει αφού οι ημιάξονές της είναι $\kappa b, \kappa c$ και το κ μειώνεται από 1 σε 0. Τα ανάλογα ισχύουν όταν το x_0 μειώνεται στο διάστημα $(-a, 0]$. Όταν $x_0 = a$ ή $x_0 = -a$, το σχήμα αποτελείται από ένα μόνο σημείο: το $(a, 0, 0)$ ή το $(-a, 0, 0)$ αντιστοίχως. Έχουμε ανάλογα συμπεράσματα για τομές παράλληλες στο xy -επίπεδο ή στο xz -επίπεδο.

3. Η καρτεσιανή εξίσωση

$$x^2 + y^2 = r_0^2$$

(και το z ελεύθερο) ορίζει τον **κύλινδρο** η οποίος είναι συμμετρικός ως προς τον z -άξονα, είναι κάθετος στο xy -επίπεδο και τέμνει το xy -επίπεδο στον κύκλο με εξίσωση $x^2 + y^2 = r_0^2$. Ο κύλινδρος εκτείνεται απεριόριστα προς τα πάνω και προς τα κάτω. Η τομή του κυλίνδρου με οποιοδήποτε επίπεδο παράλληλο στο xy -επίπεδο (δηλαδή κάθετο στον z -άξονα) είναι κύκλος ακτίνας r_0 και με κέντρο πάνω στον z -άξονα.

Μέρος του προηγούμενου κυλίνδρου παίρνουμε αν περιορίσουμε τις τιμές του z :

$$x^2 + y^2 = r_0^2, \quad z_1 \leq z \leq z_2.$$

Τώρα έχουμε το μέρος του προηγούμενου κυλίνδρου το οποίο βρίσκεται ανάμεσα στα δύο επίπεδα που είναι παράλληλα στο xy -επίπεδο με καρτεσιανές εξισώσεις $z = z_1$ και $z = z_2$.

Η κυλινδρική εξίσωση του κυλίνδρου είναι η

$$r = r_0$$

(το θ είναι ελεύθερο στο διάστημα $[0, 2\pi]$ και το z είναι ελεύθερο).

Η κυλινδρική εξίσωση του περιορισμένου κυλίνδρου είναι η

$$r = r_0, \quad z_1 \leq z \leq z_2$$

(το θ είναι ελεύθερο στο διάστημα $[0, 2\pi]$).

Ανάλογους κυλίνδρους παίρνουμε από τις εξισώσεις $y^2 + z^2 = r_0^2$ (και το x ελεύθερο ή περιορισμένο σε διάστημα $[x_1, x_2]$) και $x^2 + z^2 = r_0^2$ (και το y ελεύθερο ή περιορισμένο σε διάστημα $[y_1, y_2]$).

4. Η καρτεσιανή εξίσωση

$$x^2 + y^2 - z^2 = 0$$

ορίζει δύο **κώνους**. Ο ένας έχει καρτεσιανή εξίσωση $z = \sqrt{x^2 + y^2}$ και ο άλλος έχει καρτεσιανή εξίσωση $z = -\sqrt{x^2 + y^2}$.

Σχετικά με τον κώνο με καρτεσιανή εξίσωση $z = \sqrt{x^2 + y^2}$ είδαμε τα εξής.

Ο κώνος έχει κορυφή το σημείο $(0, 0, 0)$, είναι συμμετρικός ως προς τον z -άξονα και ανοίγει προς τα πάνω.

Η τομή του κώνου με οποιοδήποτε επίπεδο παράλληλο στο xy -επίπεδο και πάνω από αυτό είναι κύκλος. Πιο συγκεκριμένα, η τομή του κώνου με το επίπεδο με καρτεσιανή εξίσωση $z = z_0$, όπου $z_0 > 0$, είναι κύκλος με κέντρο το σημείο $(0, 0, z_0)$ πάνω στον z -άξονα και ακτίνα z_0 . Όταν αυξάνεται το z_0 στο διάστημα $(0, +\infty)$ ο κύκλος ανεβαίνει προς τα πάνω και αυξάνεται η ακτίνα του. Η ένωση όλων αυτών των κύκλων σχηματίζει τον κώνο.

Η τομή του κώνου με οποιοδήποτε ημιεπίπεδο το οποίο έχει ως συνοριακή ευθεία τον z -άξονα είναι ημιευθεία. Πιο συγκεκριμένα, αν θεωρήσουμε στο xy -επίπεδο την ημιευθεία με κορυφή το $(0, 0, 0)$ η οποία έχει πολική γωνία θ (με τον θετικό x -ημιάξονα), τότε αυτή η ημιευθεία μαζί με τον z -άξονα σχηματίζουν ένα ημιεπίπεδο στον χώρο με συνοριακή ευθεία τον z -άξονα. Τα σημεία αυτού του ημιεπιπέδου καθορίζονται από το ζεύγος (κυλινδρικών) συντεταγμένων (r, z) . Η τομή του κώνου με αυτό το ημιεπίπεδο είναι η ημιευθεία με εξίσωση $z = r$. Η ημιευθεία αυτή έχει γωνία $\phi = \frac{\pi}{4}$ με τον θετικό z -ημιάξονα. Όταν η γωνία θ (η οποία καθορίζει το ημιεπίπεδο) αυξάνεται στο διάστημα $[0, 2\pi]$ το ημιεπίπεδο περιστρέφεται γύρω από τον z -άξονα και η ημιευθεία (η τομή του κώνου με το ημιεπίπεδο) περιστρέφεται γύρω από τον z -άξονα διατηρώντας σταθερή γωνία $\phi = \frac{\pi}{4}$ με τον θετικό z -άξονα. Η ένωση όλων αυτών των ημιευθειών σχηματίζει τον κώνο.

Είδαμε ότι η κυλινδρική εξίσωση του κώνου είναι η

$$z = r$$

(και το θ ελεύθερο στο διάστημα $[0, 2\pi]$).

Η σφαιρική εξίσωση του κώνου είναι η

$$\phi = \frac{\pi}{4}$$

(το ρ ελεύθερο στο διάστημα $[0, +\infty)$ και το θ ελεύθερο στο διάστημα $[0, 2\pi]$).

Ο κώνος με καρτεσιανή εξίσωση $z = -\sqrt{x^2 + y^2}$ είναι συμμετρικός του προηγούμενου ως προς το xy -επίπεδο. Έχει κορυφή το σημείο $(0, 0, 0)$, είναι συμμετρικός ως προς τον z -άξονα και ανοίγει προς τα κάτω.

Η κυλινδρική εξίσωση του κώνου είναι η

$$z = -r$$

(και το θ ελεύθερο στο διάστημα $[0, 2\pi]$) και η σφαιρική εξίσωσή του είναι η

$$\phi = \frac{3\pi}{4}$$

(το ρ ελεύθερο στο διάστημα $[0, +\infty)$ και το θ ελεύθερο στο διάστημα $[0, 2\pi]$).

Εντελώς ανάλογες περιγραφές έχουμε για τους κώνους με καρτεσιανές εξισώσεις

$$-x^2 + y^2 + z^2 = 0 \quad \text{και} \quad x^2 - y^2 + z^2 = 0.$$

(Αυτοί δεν έχουν απλές κυλινδρικές και σφαιρικές εξισώσεις.)

5. Η καρτεσιανή εξίσωση

$$z = x^2 + y^2$$

ορίζει ένα **παραβολοειδές εκ περιστροφής**.

Η επιφάνεια αυτή έχει κορυφή το σημείο $(0, 0, 0)$, είναι συμμετρική ως προς τον z -άξονα και ανοίγει προς τα πάνω.

Η τομή της επιφάνειας με οποιοδήποτε επίπεδο παράλληλο στο xy -επίπεδο και πάνω από αυτό είναι κύκλος. Πιο συγκεκριμένα, η τομή της επιφάνειας με το επίπεδο με καρτεσιανή εξίσωση $z = z_0$, όπου $z_0 > 0$, είναι κύκλος με κέντρο το σημείο $(0, 0, z_0)$ πάνω στον z -άξονα και ακτίνα $\sqrt{z_0}$. Όταν αυξάνεται το z_0 στο διάστημα $(0, +\infty)$ ο κύκλος ανεβαίνει προς τα πάνω και αυξάνεται η ακτίνα του. Η ένωση όλων αυτών των κύκλων σχηματίζει την επιφάνεια.

Η τομή της επιφάνειας με οποιοδήποτε ημιεπίπεδο το οποίο έχει ως συνοριακή ευθεία τον z -άξονα είναι μισή παραβολή. Πιο συγκεκριμένα, αν θεωρήσουμε στο xy -επίπεδο την ημιευθεία με κορυφή το $(0, 0, 0)$ η οποία έχει πολική γωνία θ (με τον θετικό x -ημιάξονα), τότε αυτή η ημιευθεία μαζί με τον z -άξονα σχηματίζουν ένα ημιεπίπεδο στον χώρο με συνοριακή ευθεία τον z -άξονα. Τα σημεία αυτού του ημιεπιπέδου καθορίζονται από το ζεύγος (κυλινδρικών) συντεταγμένων (r, z) . Η τομή της επιφάνειας με αυτό το ημιεπίπεδο είναι η μισή παραβολή με εξίσωση $z = r^2$. Όταν η γωνία θ (η οποία καθορίζει το ημιεπίπεδο) αυξάνεται στο διάστημα $[0, 2\pi]$ το ημιεπίπεδο περιστρέφεται γύρω από τον z -άξονα και η μισή παραβολή (η τομή της επιφάνειας με το ημιεπίπεδο) περιστρέφεται γύρω από τον z -άξονα. Η ένωση όλων αυτών των ημι-παραβολών σχηματίζει το παραβολοειδές εκ περιστροφής.

Είδαμε ότι η κυλινδρική εξίσωση του παραβολοειδούς εκ περιστροφής είναι η

$$z = r^2$$

(και το θ ελεύθερο στο διάστημα $[0, 2\pi]$) και ότι η σφαιρική εξίσωσή του είναι η

$$\rho(\cos \phi - \rho \sin^2 \phi) = 0$$

(και το θ ελεύθερο στο διάστημα $[0, 2\pi]$).

6. Η καρτεσιανή εξίσωση

$$z = x^2 - y^2$$

όπου το c είναι δοσμένος αριθμός, ορίζει ένα **υπερβολικό παραβολοειδές**.

Η τομή αυτής της επιφάνειας με το xy -επίπεδο αποτελείται από την ένωση των κύριων διαγωνίων του xy -επιπέδου.

Η τομή της επιφάνειας με οποιοδήποτε επίπεδο παράλληλο στο xy -επίπεδο και πάνω ή κάτω από αυτό είναι δύο κλάδοι υπερβολής. Πιο συγκεκριμένα, η τομή της επιφάνειας με το επίπεδο με καρτεσιανή εξίσωση $z = z_0$, όπου $z_0 > 0$, είναι οι δύο κλάδοι της υπερβολής με καρτεσιανή εξίσωση $x^2 - y^2 = z_0$ (στο xy -επίπεδο) υπερυψωμένοι σε ύψος z_0 πάνω από το xy -επίπεδο. Ομοίως, η τομή της επιφάνειας με το επίπεδο με καρτεσιανή εξίσωση $z = z_0$, όπου $z_0 < 0$, είναι οι δύο κλάδοι της υπερβολής με καρτεσιανή εξίσωση $x^2 - y^2 = z_0$ (στο xy -επίπεδο) βυθισμένοι σε βάθος $|z_0| = -z_0$ κάτω από το xy -επίπεδο. Η ένωση όλων αυτών των υπερβολών σχηματίζει το υπερβολικό παραβολοειδές.

Η μελέτη του υπερβολικού παραβολοειδούς θα συνεχιστεί την επόμενη εβδομάδα.

4η εβδομάδα.

Είδαμε ότι οι τομές του υπερβολικού παραβολοειδούς με επίπεδα παράλληλα στο yz -επίπεδο είναι παραβολές οι οποίες ανοίγουν προς τα κάτω και οι τομές του με επίπεδα παράλληλα στο xz -επίπεδο είναι παραβολές οι οποίες ανοίγουν προς τα πάνω. Οι τομές του υπερβολικού παραβολοειδούς με επίπεδα παράλληλα στο xy -επίπεδο είναι υπερβολές (εκτός από την τομή με το ίδιο το xy -επίπεδο η οποία αποτελείται από δύο ευθείες).

Αρχίσαμε να μιλάμε για συναρτήσεις

$$f : A \rightarrow \mathbb{R}^m, \quad A \subseteq \mathbb{R}^n.$$

Η ανεξάρτητη μεταβλητή $\mathbf{x} = (x_1, \dots, x_n)$ βρίσκεται στο πεδίο ορισμού A , το οποίο είναι κάποιο υποσύνολο του $\mathbb{R}^n$, και η εξαρτημένη μεταβλητή $\mathbf{y} = (y_1, \dots, y_m)$ βρίσκεται στον $\mathbb{R}^m$. Η ανεξάρτητη και η εξαρτημένη μεταβλητή συνδέονται με τη σχέση

$$\mathbf{y} = f(\mathbf{x})$$

ή, ισοδύναμα,

$$(y_1, \dots, y_m) = f(x_1, \dots, x_n).$$

Καθεμία από τις συντεταγμένες $y_1, \dots, y_m$ της εξαρτημένης μεταβλητής $\mathbf{y}$ εξαρτάται από την ανεξάρτητη μεταβλητή $\mathbf{x}$, δηλαδή είναι συνάρτηση του $\mathbf{x}$. Έτσι μπορούμε να γράψουμε

$$\begin{aligned} y_1 &= f_1(\mathbf{x}) = f_1(x_1, \dots, x_n) \\ &\vdots \\ y_m &= f_m(\mathbf{x}) = f_m(x_1, \dots, x_n) \end{aligned}$$

όπου οι $f_1, \dots, f_m$ είναι συναρτήσεις ορισμένες στο A και με τιμές στο $\mathbb{R}$:

$$f_1, \dots, f_m : A \rightarrow \mathbb{R}, \quad A \subseteq \mathbb{R}^n.$$

Οι συναρτήσεις $f_1, \dots, f_m$ ονομάζονται **συντεταγμένες συναρτήσεις** της f .

Για παράδειγμα, η συνάρτηση $f : \mathbb{R} \rightarrow \mathbb{R}^3$ με τύπο

$$f(x) = (2x, x^2 + \sin x, e^{3x})$$

έχει τρεις συντεταγμένες συναρτήσεις $f_1, f_2, f_3 : \mathbb{R} \rightarrow \mathbb{R}$ με τύπους

$$f_1(x) = 2x, \quad f_2(x) = x^2 + \sin x, \quad f_3(x) = e^{3x}.$$

Η συνάρτηση $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ με τύπο

$$f(x, y, z) = (x^2y + ye^z, x + y \sin z)$$

έχει δύο συντεταγμένες συναρτήσεις $f_1, f_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$ με τύπους

$$f_1(x, y, z) = x^2y + ye^z, \quad f_2(x, y, z) = x + y \sin z.$$

Η συνάρτηση $f : A \rightarrow \mathbb{R}$ με τύπο

$$f(x, y) = \frac{1}{xy}$$

έχει πεδίο ορισμού το σύνολο

$$A = \{(x, y) \mid xy \neq 0\} = \{(x, y) \mid x \neq 0, y \neq 0\},$$

δηλαδή το $\mathbb{R}^2$ εκτός από την ένωση του x -άξονα και του y -άξονα. Η συνάρτηση f έχει μόνο μία συντεταγμένη συνάρτηση, τον εαυτό της.

Αν έχουμε μια συνάρτηση

$$f : A \rightarrow \mathbb{R}^m, \quad A \subseteq \mathbb{R}^n,$$

τότε στην περίπτωση $m = 1$ λέμε ότι η f είναι **πραγματική** ή **βαθμωτή** συνάρτηση ενώ στην περίπτωση $m \geq 2$ λέμε ότι η f είναι **διανυσματική** συνάρτηση. Επομένως οι συντεταγμένες συναρτήσεις μιας διανυσματικής συνάρτησης είναι πραγματικές συναρτήσεις.

Στην περίπτωση **πραγματικής** συνάρτησης (δηλαδή όταν $m = 1$)

$$f : A \rightarrow \mathbb{R}, \quad A \subseteq \mathbb{R}^n,$$

ορίσαμε το **γράφημα** της f ως το σύνολο

$$G_f = \{(\mathbf{x}, f(\mathbf{x})) \mid \mathbf{x} \in A\} = \{(x_1, \dots, x_n, f(x_1, \dots, x_n)) \mid (x_1, \dots, x_n) \in A\}.$$

Το G_f είναι υποσύνολο του $\mathbb{R}^{n+1}$ αφού κάθε

$$(\mathbf{x}, f(\mathbf{x})) = (x_1, \dots, x_n, f(x_1, \dots, x_n))$$

έχει $n + 1$ πραγματικές συντεταγμένες.

Πάλι στην περίπτωση **πραγματικής** συνάρτησης (δηλαδή όταν $m = 1$)

$$f : A \rightarrow \mathbb{R}, \quad A \subseteq \mathbb{R}^n,$$

ορίσαμε τα **ισοσταθμικά σύνολα** της f ως εξής. Για κάθε πραγματική τιμή $c \in \mathbb{R}$ το αντίστοιχο ισοσταθμικό σύνολο είναι το

$$A_c = \{\mathbf{x} \in A \mid f(\mathbf{x}) = c\} = \{(x_1, \dots, x_n) \in A \mid f(x_1, \dots, x_n) = c\}.$$

Δηλαδή κάθε A_c είναι υποσύνολο του πεδίου ορισμού $A \subseteq \mathbb{R}^n$ και αποτελείται από όλα τα σημεία του A στα οποία η f έχει σταθερή τιμή c :

$$(x_1, \dots, x_n) \in A_c \iff (x_1, \dots, x_n) \in A \text{ και } f(x_1, \dots, x_n) = c.$$

Αν η τιμή c δεν ανήκει στο σύνολο τιμών της f , τότε το αντίστοιχο ισοσταθμικό σύνολο A_c είναι κενό. Αν η τιμή c ανήκει στο σύνολο τιμών της f , τότε το αντίστοιχο ισοσταθμικό σύνολο A_c έχει τουλάχιστον ένα στοιχείο. Σε διαφορετικές τιμές c_1, c_2 του c στο σύνολο τιμών της f αντιστοιχούν ξένα ανά δύο ισοσταθμικά σύνολα A_{c_1}, A_{c_2} . Μεταβάλλοντας το c ώστε να διατρέχει το σύνολο τιμών της f παίρνουμε ολόκληρο το πεδίο ορισμού A ως την ένωση όλων των ισοσταθμικών συνόλων A_c .

Αν $n = 1$, το γράφημα μιας πραγματικής συνάρτησης είναι *συνήθως* μια καμπύλη στον $\mathbb{R}^2$. Αν $n = 2$, το γράφημα μιας πραγματικής συνάρτησης είναι *συνήθως* μια επιφάνεια στον $\mathbb{R}^3$. Σ' αυτές τις δύο περιπτώσεις υπάρχει πολλές φορές η δυνατότητα να σχεδιαστεί το γράφημα της συνάρτησης, ειδικά αν αυτή είναι σχετικά απλή. Αν, όμως, $n \geq 3$, το γράφημα είναι υποσύνολο του $\mathbb{R}^{n+1}$, όπου $n + 1 \geq 4$, και δεν μπορούμε να το σχεδιάσουμε.

Αν $n = 1$, ένα οποιοδήποτε ισοσταθμικό σύνολο μιας πραγματικής συνάρτησης αποτελείται *συνήθως* από πεπερασμένου πλήθους σημεία. Αν $n = 2$, ένα οποιοδήποτε ισοσταθμικό σύνολο

είναι συνήθως μια καμπύλη στον $\mathbb{R}^2$. Αν $n = 3$, ένα οποιοδήποτε ισοσταθμικό σύνολο είναι συνήθως μια επιφάνεια στον $\mathbb{R}^3$. Σ' αυτές τις περιπτώσεις μπορούμε να σχεδιάσουμε τα ισοσταθμικά σύνολα της συνάρτησης αν αυτή είναι σχετικά απλή. Αν, όμως, $n \geq 4$, επειδή τα ισοσταθμικά σύνολα είναι υποσύνολα του $\mathbb{R}^n$, δεν μπορούμε να τα σχεδιάσουμε.

Στην περίπτωση $n = 2$, όταν έχουμε πραγματική συνάρτηση

$$z = f(x, y)$$

είδαμε και με παραδείγματα τη γεωμετρική εικόνα του γραφήματος και των ισοσταθμικών καμπυλών της συνάρτησης αλλά και τη γεωμετρική σχέση ανάμεσα στις ισοσταθμικές καμπύλες και στο γράφημα της συνάρτησης. Το πεδίο ορισμού A της f βρίσκεται στο xy -επίπεδο και το γράφημά της, το G_f , είναι συνήθως μια επιφάνεια στον χώρο και "πάνω" από το A (πάνω, όπου οι τιμές z της f είναι ≥ 0 , και κάτω, όπου οι τιμές z είναι ≤ 0). Κάθε ισοσταθμική καμπύλη A_c περιέχεται στο A , δηλαδή είναι στο xy -επίπεδο, και ακριβώς από "πάνω" της, στον χώρο, και σε ύψος $z = c$ από το xy -επίπεδο βρίσκεται μια παράλληλη καμπύλη η οποία περιέχεται στο γράφημα της f . Αν πάρουμε την ένωση όλων αυτών των παράλληλων καμπυλών των ισοσταθμικών καμπυλών θα σχηματιστεί η επιφάνεια/γράφημα της f .

Παράδειγμα. Το γράφημα G_f της πραγματικής συνάρτησης με τύπο

$$f(x, y) = ax + by$$

είναι το επίπεδο στον $\mathbb{R}^3$ με καρτεσιανή εξίσωση $z = ax + by$ ή, ισοδύναμα, $ax + by - z = 0$, δηλαδή είναι το επίπεδο που περιέχει το σημείο $(0, 0, 0)$ και είναι κάθετο στο διάνυσμα $(a, b, -1)$. Για κάθε $c \in \mathbb{R}$ το ισοσταθμικό σύνολο A_c της f είναι η ευθεία στο xy -επίπεδο με καρτεσιανή εξίσωση $ax + by = c$. Στον χώρο, ακριβώς "πάνω" από αυτήν την ευθεία και σε ύψος $z = c$ από το xy -επίπεδο βρίσκεται μια παράλληλη ευθεία η οποία περιέχεται στο επίπεδο/γράφημα της συνάρτησης. Η ένωση όλων (δηλαδή για όλες τις τιμές του c) αυτών των παράλληλων ευθειών σχηματίζει το επίπεδο/γράφημα της συνάρτησης.

Παράδειγμα. Το γράφημα G_f της πραγματικής συνάρτησης με τύπο

$$z = f(x, y) = x^2 + y^2$$

είναι το παραβολοειδές εκ περιστροφής στον $\mathbb{R}^3$ με καρτεσιανή εξίσωση $z = x^2 + y^2$. Αν $c < 0$, το ισοσταθμικό σύνολο A_c της f είναι κενό. Αν $c = 0$, το ισοσταθμικό σύνολο A_c αποτελείται από ένα μόνο σημείο, το $(0, 0)$. Για κάθε $c > 0$ το ισοσταθμικό σύνολο A_c είναι ο κύκλος στο xy -επίπεδο με καρτεσιανή εξίσωση $x^2 + y^2 = c$, δηλαδή ο κύκλος με κέντρο το $(0, 0)$ και ακτίνα $\sqrt{c}$. Στον χώρο, ακριβώς πάνω από αυτόν τον κύκλο και σε ύψος $z = c > 0$ από το xy -επίπεδο βρίσκεται ένας παράλληλος κύκλος ο οποίος περιέχεται στο παραβολοειδές εκ περιστροφής. Η ένωση όλων (δηλαδή για όλες τις θετικές τιμές του c) αυτών των παράλληλων κύκλων μαζί με το σημείο $(0, 0, 0)$ σχηματίζει το παραβολοειδές εκ περιστροφής, δηλαδή το γράφημα της f .

Αν $\mathbf{x} \in \mathbb{R}^n$ και $r > 0$ η **μπάλα** με κέντρο $\mathbf{x}$ και ακτίνα r είναι το σύνολο των σημείων $\mathbf{y}$ του $\mathbb{R}^n$ σε απόσταση από το $\mathbf{x}$ μικρότερη από r . Την μπάλα αυτή συμβολίζουμε $B_r(\mathbf{x})$, οπότε

$$B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n \mid \|\mathbf{y} - \mathbf{x}\| < r\}.$$

Αν $n = 3$, τότε η έννοια της μπάλας όπως την ορίσαμε ταυτίζεται με τη γνωστή έννοια της μπάλας στο χώρο. Αν $n = 2$, τότε η έννοια της μπάλας όπως την ορίσαμε ταυτίζεται με τη γνωστή έννοια του δίσκου στο επίπεδο. Ενώ, αν $n = 1$, η έννοια της μπάλας όπως την ορίσαμε ταυτίζεται με τη γνωστή έννοια του διαστήματος στην ευθεία.

Αν $A \subseteq \mathbb{R}^n$ και $\mathbf{x} \in A$ είπαμε ότι το x είναι **εσωτερικό σημείο** του A αν υπάρχει κάποιο $\delta > 0$ (ίσως αρκετά μικρό) ώστε η μπάλα $B_\delta(\mathbf{x})$ να περιέχεται ολόκληρη στο A ή, ισοδύναμα, ώστε κάθε $\mathbf{y}$ το οποίο απέχει από το $\mathbf{x}$ λιγότερο από δ να περιέχεται στο A :

$$B_\delta(\mathbf{x}) \subseteq A \iff [\|\mathbf{y} - \mathbf{x}\| < \delta \Rightarrow \mathbf{y} \in A].$$

Αν $A \subseteq \mathbb{R}^n$ και $\mathbf{x} \in A$ είπαμε ότι το x είναι **συνοριακό σημείο** του A αν για κάθε $\delta > 0$ (οσοδήποτε μικρό) η μπάλα $B_\delta(\mathbf{x})$ τέμνει και το A αλλά και το συμπληρωματικό σύνολο του A ή, ισοδύναμα, ώστε για κάθε $\delta > 0$ υπάρχουν σημεία του A και σημεία έξω από το A τα οποία απέχουν από το $\mathbf{x}$ λιγότερο από δ .

Είπαμε ότι ένα $A \subseteq \mathbb{R}^n$ ονομάζεται **ανοικτό** αν όλα τα σημεία του είναι εσωτερικά σημεία του ή, ισοδύναμα, κανένα σημείο του δεν είναι συνοριακό σημείο του. Είπαμε ότι ένα $A \subseteq \mathbb{R}^n$ ονομάζεται **κλειστό** αν, εκτός από τα εσωτερικά σημεία του, περιέχει και όλα τα συνοριακά σημεία του:

A ανοικτό $\iff$ το A περιέχει όλα τα εσωτερικά σημεία του και κανένα συνοριακό σημείο του
 A κλειστό $\iff$ το A περιέχει όλα τα εσωτερικά σημεία του και όλα τα συνοριακά σημεία του

Παράδειγμα. Τα σημεία που βρίσκονται μέσα σε ένα δίσκο στο επίπεδο είναι τα εσωτερικά του σημεία και τα σημεία που βρίσκονται στην περιφέρεια του δίσκου είναι τα συνοριακά του σημεία. Άρα ένας δίσκος χωρίς κανένα από τα περιφερειακά του σημεία είναι ανοικτός ενώ ένας δίσκος με όλα τα περιφερειακά του σημεία είναι κλειστός:

$$\text{ανοικτός δίσκος: } \{(x, y) \mid (x - x_0)^2 + (y - y_0)^2 < r_0^2\}$$

$$\text{κλειστός δίσκος: } \{(x, y) \mid (x - x_0)^2 + (y - y_0)^2 \leq r_0^2\}$$

Ένας δίσκος ο οποίος περιέχει κάποια, αλλά όχι όλα, από τα περιφερειακά του σημεία δεν είναι ούτε ανοικτός ούτε κλειστός.

Παράδειγμα. Τα σημεία που βρίσκονται μέσα σε μια μπάλα στον χώρο είναι τα εσωτερικά της σημεία και τα σημεία που βρίσκονται στην επιφάνεια της μπάλας είναι τα συνοριακά της σημεία. Άρα μία μπάλα χωρίς κανένα από τα επιφανειακά της σημεία είναι ανοικτή ενώ μία μπάλα με όλα τα επιφανειακά της σημεία είναι κλειστή:

$$\text{ανοικτή μπάλα: } \{(x, y, z) \mid (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 < r_0^2\}$$

$$\text{κλειστή μπάλα: } \{(x, y, z) \mid (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 \leq r_0^2\}$$

Μια μπάλα η οποία περιέχει κάποια, αλλά όχι όλα, από τα επιφανειακά της σημεία δεν είναι ούτε ανοικτή ούτε κλειστή.

Παράδειγμα. Τα σημεία που βρίσκονται μέσα σε ένα διάστημα στην ευθεία είναι τα εσωτερικά του σημεία και τα άκρα του διαστήματος είναι τα συνοριακά του σημεία. Άρα ένα διάστημα χωρίς κανένα από τα άκρα του είναι ανοικτό ενώ ένα διάστημα με τα δύο άκρα του είναι κλειστό:

$$\text{ανοικτό διάστημα: } \{x \mid |x - x_0| < r_0\}$$

$$\text{κλειστό διάστημα: } \{x \mid |x - x_0| \leq r_0\}$$

Ένα διάστημα το οποίο περιέχει μόνο ένα από τα δύο άκρα του δεν είναι ούτε ανοικτό ούτε κλειστό.

5η εβδομάδα.

Ορίσαμε την έννοια του ορίου συνάρτησης:

$$f : A \rightarrow \mathbb{R}^m, \quad A \subseteq \mathbb{R}^n,$$

όπου το σύνολο A αποτελείται από ένα ανοικτό υποσύνολο του $\mathbb{R}^n$ μαζί με κάποια (μπορεί κανένα, μπορεί όλα, μπορεί μόνο μερικά) από τα συνοριακά του σημεία. Αν το $\mathbf{x}_0$ είναι εσωτερικό σημείο ή συνοριακό σημείο του A (δεν είναι ανάγκη να ανήκει στο A) και αν $\mathbf{y}_0 \in \mathbb{R}^m$, λέμε ότι το **όριο** της f στο $\mathbf{x}_0$ είναι το $\mathbf{y}_0$ και συμβολίζουμε

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0$$

αν

για κάθε $\epsilon > 0$ υπάρχει $\delta > 0$ ώστε : $\mathbf{x} \in A, 0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta \Rightarrow \|f(\mathbf{x}) - \mathbf{y}_0\| < \epsilon$.

Σε απλοϊκή γλώσσα: αν το $\mathbf{x}$ ανήκει στο πεδίο ορισμού της συνάρτησης f και πλησιάζει το $\mathbf{x}_0$, παραμένοντας διαφορετικό από το $\mathbf{x}_0$, τότε το $\mathbf{y} = f(\mathbf{x})$ (το οποίο είναι στον $\mathbb{R}^m$) πλησιάζει το $\mathbf{y}_0$.

Στην περίπτωση $m = 1$, δηλαδή όταν η f είναι πραγματική, λέμε ότι το όριο της f στο $\mathbf{x}_0$ είναι το $+\infty$ και συμβολίζουμε

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = +\infty$$

αν

για κάθε $M > 0$ υπάρχει $\delta > 0$ ώστε : $\mathbf{x} \in A, 0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta \Rightarrow f(\mathbf{x}) > M$.

Σε απλοϊκή γλώσσα: αν το $\mathbf{x}$ ανήκει στο πεδίο ορισμού της συνάρτησης f και πλησιάζει το $\mathbf{x}_0$, παραμένοντας διαφορετικό από το $\mathbf{x}_0$, τότε το $y = f(\mathbf{x})$ (το οποίο είναι στον $\mathbb{R}$) απομακρύνεται προς το $+\infty$.

Ομοίως, πάλι στην περίπτωση $m = 1$, λέμε ότι το όριο της f στο $\mathbf{x}_0$ είναι το $-\infty$ και συμβολίζουμε

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = -\infty$$

αν

για κάθε $M > 0$ υπάρχει $\delta > 0$ ώστε : $\mathbf{x} \in A, 0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta \Rightarrow f(\mathbf{x}) < -M$.

Σε απλοϊκή γλώσσα: αν το $\mathbf{x}$ ανήκει στο πεδίο ορισμού της συνάρτησης f και πλησιάζει το $\mathbf{x}_0$, παραμένοντας διαφορετικό από το $\mathbf{x}_0$, τότε το $y = f(\mathbf{x})$ (το οποίο είναι στον $\mathbb{R}$) απομακρύνεται προς το $-\infty$.

Αν το $\mathbf{x}_0$ ανήκει στο πεδίο ορισμού A της f και αν $\mathbf{y}_0 = f(\mathbf{x}_0)$, τότε έχουμε την έννοια της συνέχειας. Λέμε ότι η f είναι **συνεχής** στο $\mathbf{x}_0$ αν

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0),$$

δηλαδή αν

για κάθε $\epsilon > 0$ υπάρχει $\delta > 0$ ώστε : $\mathbf{x} \in A, \|\mathbf{x} - \mathbf{x}_0\| < \delta \Rightarrow \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \epsilon$.

Σε απλοϊκή γλώσσα: αν το $\mathbf{x}$ ανήκει στο πεδίο ορισμού της συνάρτησης f και πλησιάζει το $\mathbf{x}_0$ τότε το $\mathbf{y} = f(\mathbf{x})$ (το οποίο είναι στον $\mathbb{R}^m$) πλησιάζει το $f(\mathbf{x}_0)$.

Αναφέραμε διάφορες ιδιότητες τον ορίων και της συνέχειας.

1. Αν $f, g : A \rightarrow \mathbb{R}^m$ και υπάρχουν τα όρια $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}), \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x})$ τότε υπάρχει και το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (f(\mathbf{x}) \pm g(\mathbf{x}))$ και

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (f(\mathbf{x}) \pm g(\mathbf{x})) = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \pm \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}).$$

(Αυτήν την ιδιότητα την αποδείξαμε.)

Αναλόγως, αν οι f, g είναι συνεχείς στο $\mathbf{x}_0$ τότε και οι $f \pm g$ είναι συνεχείς στο $\mathbf{x}_0$.

2. Αν $f : A \rightarrow \mathbb{R}^m$ και υπάρχει το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ και αν $\lambda \in \mathbb{R}$ τότε υπάρχει και το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \lambda f(\mathbf{x})$ και

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \lambda f(\mathbf{x}) = \lambda \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}).$$

Αναλόγως, αν η f είναι συνεχής στο $\mathbf{x}_0$ και αν $\lambda \in \mathbb{R}$ τότε και η λf είναι συνεχής στο $\mathbf{x}_0$.

3. Αν $f, g : A \rightarrow \mathbb{R}$ (προσέξτε: $m = 1$) και υπάρχουν τα όρια $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}), \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x})$ τότε υπάρχει και το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})g(\mathbf{x})$ και

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})g(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}).$$

Αναλόγως, αν οι f, g είναι συνεχείς στο $\mathbf{x}_0$ τότε και η fg είναι συνεχής στο $\mathbf{x}_0$.

4. Αν $f, g : A \rightarrow \mathbb{R}$ (προσέξτε: $m = 1$) και υπάρχουν τα όρια $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}), \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x})$ και αν $g(\mathbf{x}) \neq 0$ για κάθε $\mathbf{x} \in A$ και $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}) \neq 0$ τότε υπάρχει και το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (f(\mathbf{x})/g(\mathbf{x}))$ και

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (f(\mathbf{x})/g(\mathbf{x})) = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) / \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}).$$

Αναλόγως, αν οι f, g είναι συνεχείς στο $\mathbf{x}_0$ και $g(\mathbf{x}) \neq 0$ για κάθε $\mathbf{x} \in A$ τότε και η f/g είναι συνεχής στο $\mathbf{x}_0$.

5. Αν $f, g : A \rightarrow \mathbb{R}^m$ και υπάρχουν τα όρια $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}), \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x})$ τότε υπάρχει και το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \cdot g(\mathbf{x})$ και

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \cdot g(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \cdot \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}).$$

Αναλόγως, αν οι f, g είναι συνεχείς στο $\mathbf{x}_0$ τότε και η $f \cdot g$ είναι συνεχής στο $\mathbf{x}_0$.

6. Αν $f, g : A \rightarrow \mathbb{R}^3$ (προσέξτε: $m = 3$) και υπάρχουν τα όρια $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}), \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x})$ τότε υπάρχει και το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \times g(\mathbf{x})$ και

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \times g(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) \times \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(\mathbf{x}).$$

Αναλόγως, αν οι f, g είναι συνεχείς στο $\mathbf{x}_0$ τότε και η $f \times g$ είναι συνεχής στο $\mathbf{x}_0$.

7. Αν $f : A \rightarrow B$, όπου $A \subseteq \mathbb{R}^n$ και $B \subseteq \mathbb{R}^m$, και $g : B \rightarrow \mathbb{R}^k$ και υπάρχει το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0$ και το όριο $\lim_{\mathbf{y} \rightarrow \mathbf{y}_0} g(\mathbf{y})$ και αν $f(\mathbf{x}) \neq \mathbf{y}_0$ για $\mathbf{x}$ κοντά στο $\mathbf{x}_0$ τότε υπάρχει και το όριο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(f(\mathbf{x}))$ και

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} g(f(\mathbf{x})) = \lim_{\mathbf{y} \rightarrow \mathbf{y}_0} g(\mathbf{y}).$$

Αναλόγως, αν η f είναι συνεχής στο $\mathbf{x}_0$ και η g είναι συνεχής στο $\mathbf{y}_0 = f(\mathbf{x}_0)$ τότε και η $g \circ f$ είναι συνεχής στο $\mathbf{x}_0$.

8. Έστω $f : A \rightarrow \mathbb{R}^m$ και έστω $f_1, \dots, f_m : A \rightarrow \mathbb{R}$ οι συντεταγμένες συναρτήσεις της f . Επίσης έστω $\mathbf{y}_0 = (y_{01}, \dots, y_{0m})$. Τότε

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0 \quad \Leftrightarrow \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_i(\mathbf{x}) = y_{0i}, \quad i = 1, \dots, m.$$

Αναλόγως, η f είναι συνεχής στο $\mathbf{x}_0$ αν και μόνο αν κάθε f_i είναι συνεχής στο $\mathbf{x}_0$.

Αφού ξεκαθαρίσαμε ότι

$$\mathbf{x} \rightarrow \mathbf{x}_0 \quad \Leftrightarrow \quad x_i \rightarrow x_{0i}, \quad i = 1, \dots, n,$$

όπου $\mathbf{x} = (x_1, \dots, x_n)$ και $\mathbf{x}_0 = (x_{01}, \dots, x_{0n})$, είδαμε διάφορα παραδείγματα ορίων και συνέχειας.

1. Κάθε συνάρτηση οι συντεταγμένες συναρτήσεις της οποίας είναι απλοί συνδυασμοί πολυωνυμικών, ρητών, εκθετικών, τριγωνομετρικών συναρτήσεων των συντεταγμένων της ανεξάρτητης μεταβλητής $\mathbf{x} = (x_1, \dots, x_n)$ είναι συνεχής. Για παράδειγμα, οι συναρτήσεις

$$f(x, y) = xy, \quad f(x, y) = (xe^{x+y}, \sin(x^2 + y)), \quad f(x, y) = (x^2, xy, e^{\sin(xy)}),$$

$$f(x, y) = (1/(xy), x^2 + y), \quad f(x, y, z) = (x/y, x^2 \sin(e^y)), \quad f(x, y, z) = 1/(x^2 + y^2 - z^2)$$

είναι όλες συνεχείς συναρτήσεις (στα αντίστοιχα πεδία ορισμού τους).

2. Είδαμε ότι

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\sin(x^2 + y^2)}{x^2 + y^2} = 1$$

με δύο τρόπους.

Πρώτος τρόπος. Χρησιμοποιήσαμε το ότι $\lim_{(x,y) \rightarrow (0,0)} (x^2 + y^2) = 0$ και το ότι $\lim_{t \rightarrow 0} \frac{\sin t}{t} = 1$ και τον κανόνα ορίου σύνθετης συνάρτησης.

Δεύτερος τρόπος. Χρησιμοποιήσαμε τους τύπους αλλαγής σε πολικές συντεταγμένες

$$x = r \cos \theta, \quad y = r \sin \theta$$

και γράψαμε

$$\frac{\sin(x^2 + y^2)}{x^2 + y^2} = \frac{\sin(r^2)}{r^2}$$

και είδαμε ότι

$$(x, y) \rightarrow (0, 0) \quad \Leftrightarrow \quad r \rightarrow 0+.$$

Μελετήσαμε μερικές τεχνικές εύρεσης ορίου ή απόδειξης μη-ύπαρξης ορίου.

Πρώτη τεχνική. Αν θέλουμε να αποδείξουμε ότι $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{0}$, γράφουμε

$$0 \leq \|f(\mathbf{x})\| \leq \dots \leq g(\mathbf{x}),$$

όπου ξεκινάμε από την παράσταση $\|f(\mathbf{x})\|$ και, εφαρμόζοντας κατάλληλες ανισότητες προς τα πάνω, φτάνουμε κάποια στιγμή σε μια απλούστερη παράσταση $g(\mathbf{x})$ η οποία εμφανώς έχει όριο 0 στο $\mathbf{x}_0$. Κατόπιν εφαρμόζουμε το γνωστό κριτήριο παρεμβολής.

Αν θέλουμε να αποδείξουμε ότι $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0$, εφαρμόζουμε το προηγούμενο για να δείξουμε το ισοδύναμο $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} (f(\mathbf{x}) - \mathbf{y}_0) = \mathbf{0}$ γράφοντας

$$0 \leq \|f(\mathbf{x}) - \mathbf{y}_0\| \leq \dots \leq g(\mathbf{x})$$

κλπ.

Δεύτερη τεχνική. Μπορούμε να κάνουμε αλλαγή σε πολικές συντεταγμένες, αν είμαστε στον $\mathbb{R}^2$, ή σε σφαιρικές συντεταγμένες, αν είμαστε στον $\mathbb{R}^3$, με τους γνωστούς τύπους:

$$x = r \cos \theta, \quad y = r \sin \theta$$

$$x = \rho \cos \theta \sin \phi, \quad y = \rho \sin \theta \sin \phi, \quad z = \rho \cos \phi.$$

Χρησιμοποιούμε τις ισοδυναμίες:

$$(x, y) \rightarrow (0, 0) \quad \Leftrightarrow \quad r \rightarrow 0+$$

$$(x, y, z) \rightarrow (0, 0) \quad \Leftrightarrow \quad \rho \rightarrow 0+$$

οι οποίες προκύπτουν από τις σχέσεις: $r = \sqrt{x^2 + y^2}$ και $\rho = \sqrt{x^2 + y^2 + z^2}$.

Τρίτη τεχνική. Το ότι ισχύει $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0$ σημαίνει ότι το $f(\mathbf{x})$ τείνει στο $\mathbf{y}_0$ όταν το $\mathbf{x}$ πλησιάζει στο $\mathbf{x}_0$ **ανεξάρτητα** από τον τρόπο με τον οποίο το $\mathbf{x}$ πλησιάζει στο $\mathbf{x}_0$.

Υπάρχουν περιπτώσεις όπου συμβαίνει το εξής. Όταν το $\mathbf{x}$ πλησιάζει το $\mathbf{x}_0$ από μια συγκεκριμένη κατεύθυνση (π.χ. κινούμενο πάνω σε μια συγκεκριμένη ημιευθεία με κορυφή το $\mathbf{x}_0$) το αντίστοιχο $f(\mathbf{x})$ τείνει σε κάποιο όριο ενώ όταν το $\mathbf{x}$ πλησιάζει το $\mathbf{x}_0$ από μια συγκεκριμένη κατεύθυνση διαφορετική από την προηγούμενη (π.χ. κινούμενο πάνω σε μια συγκεκριμένη ημιευθεία με κορυφή το $\mathbf{x}_0$ διαφορετική από την προηγούμενη ημιευθεία) το αντίστοιχο $f(\mathbf{x})$ τείνει σε κάποιο όριο διαφορετικό από το προηγούμενο όριο. Τότε το συμπέρασμα είναι ότι το $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ **δεν** υπάρχει.

3. Έχουμε ότι

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2}{\sqrt{x^2 + y^2}} = 0.$$

Αυτό προκύπτει από την ανισότητα

$$0 \leq \left| \frac{x^2}{\sqrt{x^2 + y^2}} \right| = \frac{|x| |x|}{\sqrt{x^2 + y^2}} \leq |x|$$

διότι $|x| \rightarrow 0+$ όταν $(x, y) \rightarrow (0, 0)$.

Επίσης προκύπτει με αλλαγή σε πολικές συντεταγμένες:

$$\frac{x^2}{\sqrt{x^2 + y^2}} = \frac{r^2 \cos^2 \theta}{r} = r \cos^2 \theta \rightarrow 0,$$

διότι $r \rightarrow 0+$ και το $\cos^2 \theta$ είναι φραγμένο.

4. Έχουμε ότι

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y}{x^2 + y^2} = 0.$$

Αυτό προκύπτει από την ανισότητα

$$0 \leq \left| \frac{x^2 y}{x^2 + y^2} \right| = \frac{x^2 |y|}{x^2 + y^2} \leq |y|.$$

Επίσης, με αλλαγή σε πολικές συντεταγμένες:

$$\frac{x^2 y}{x^2 + y^2} = \frac{r^3 \cos^2 \theta \sin \theta}{r^2} = r \cos^2 \theta \sin \theta \rightarrow 0.$$

5. Έχουμε ότι

$$\lim_{(x,y,z) \rightarrow (0,0,0)} \frac{xyz}{x^2 + y^2 + z^2} = 0.$$

Αυτό προκύπτει από την ανισότητα

$$0 \leq \left| \frac{xyz}{x^2 + y^2 + z^2} \right| = \frac{|x||y||z|}{x^2 + y^2 + z^2} \leq \frac{(\sqrt{x^2 + y^2 + z^2})^3}{x^2 + y^2 + z^2} = \sqrt{x^2 + y^2 + z^2}.$$

Επίσης, με αλλαγή σε σφαιρικές συντεταγμένες:

$$\frac{xyz}{x^2 + y^2 + z^2} = \frac{\rho^3 \cos \theta \sin \theta \sin^2 \phi \cos \phi}{\rho^2} = \rho \cos \theta \sin \theta \sin^2 \phi \cos \phi \rightarrow 0,$$

διότι $\rho \rightarrow 0+$ και το $\cos \theta \sin \theta \sin^2 \phi \cos \phi$ είναι φραγμένο.

6. Το

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2}{x^2 + y^2}$$

δεν υπάρχει. Η συνάρτηση $f(x, y) = \frac{x^2}{x^2 + y^2}$ τείνει σε διαφορετικά όρια όταν το (x, y) πλησιάζει στο $(0, 0)$ κινούμενο πάνω στον x -άξονα και κινούμενο πάνω στον y -άξονα. Όταν το (x, y) κινείται πάνω στον x -άξονα έχουμε $y = 0$ και $x \rightarrow 0$, οπότε

$$f(x, 0) = \frac{x^2}{x^2 + 0^2} = 1 \rightarrow 1.$$

Όταν το (x, y) κινείται πάνω στον y -άξονα έχουμε $x = 0$ και $y \rightarrow 0$, οπότε

$$f(0, y) = \frac{0^2}{0^2 + y^2} = 0 \rightarrow 0.$$

7. Το

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2}$$

δεν υπάρχει. Δοκιμάζοντας τις τιμές της συνάρτησης $f(x, y) = \frac{xy}{x^2 + y^2}$ όταν το (x, y) πλησιάζει στο $(0, 0)$ κινούμενο πάνω στον x -άξονα και κινούμενο πάνω στον y -άξονα βρίσκουμε τα ίδια όρια. Όταν το (x, y) κινείται πάνω στον x -άξονα έχουμε $y = 0$ και $x \rightarrow 0$, οπότε

$$f(x, 0) = \frac{x0}{x^2 + 0^2} = 0 \rightarrow 0.$$

Όταν το (x, y) κινείται πάνω στον y -άξονα έχουμε $x = 0$ και $y \rightarrow 0$, οπότε

$$f(0, y) = \frac{0y}{0^2 + y^2} = 0 \rightarrow 0.$$

Όμως, όταν το (x, y) κινείται πάνω στην κύρια διαγώνιο με καρτεσιανή εξίσωση $y = x$ έχουμε $(x, y) = (x, x)$ και $x \rightarrow 0$, οπότε

$$f(x, x) = \frac{x^2}{x^2 + x^2} = \frac{1}{2} \rightarrow \frac{1}{2}.$$

Μπορούμε, επίσης, να πάρουμε περισσότερη πληροφορία δοκιμάζοντας τις τιμές της f όταν το (x, y) πλησιάζει το $(0, 0)$ κινούμενο πάνω σε ημιευθεία η οποία σχηματίζει πολική γωνία θ με τον θετικό x -άξονα. (Η περίπτωση του x -άξονα είναι ειδική περίπτωση με $\theta = 0$ και $\theta = \pi$ και η περίπτωση του y -άξονα είναι ειδική περίπτωση με $\theta = \frac{\pi}{2}$ και $\theta = \frac{3\pi}{2}$.) Αυτομάτως χρησιμοποιούμε πολικές συντεταγμένες: όταν το (x, y) πλησιάζει το $(0, 0)$ κινούμενο πάνω σε ημιευθεία η οποία σχηματίζει πολική γωνία θ με τον θετικό x -άξονα τότε έχουμε ότι

$$x = r \cos \theta, \quad y = r \sin \theta$$

και $r \rightarrow 0+$ ενώ το θ είναι συγκεκριμένο και σταθερό. Τότε

$$f(x, y) = f(r \cos \theta, r \sin \theta) = \frac{r^2 \cos \theta \sin \theta}{r^2} = \cos \theta \sin \theta \rightarrow \cos \theta \sin \theta.$$

Άρα η τιμή του ορίου εξαρτάται από την γωνία θ και έχει διαφορετικές τιμές όταν αλλάζουμε την ημιευθεία προσέγγισης του (x, y) στο $(0, 0)$.

8. Το

$$\lim_{(x,y,z) \rightarrow (0,0,0)} \frac{xyz}{(x^2 + y^2 + z^2)\sqrt{x^2 + y^2 + z^2}}$$

δεν υπάρχει. Δοκιμάζοντας τις τιμές της συνάρτησης $f(x, y, z) = \frac{xyz}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}$ όταν το (x, y, z) πλησιάζει στο $(0, 0, 0)$ κινούμενο πάνω σε ημιευθεία η οποία σχηματίζει σφαιρικές γωνίες θ, ϕ , χρησιμοποιούμε σφαιρικές συντεταγμένες: έχουμε ότι

$$x = \rho \cos \theta \sin \phi, \quad y = \rho \sin \theta \sin \phi, \quad z = \rho \cos \phi$$

και $\rho \rightarrow 0+$ ενώ τα θ, ϕ είναι συγκεκριμένα και σταθερά. Τότε

$$\begin{aligned} f(x, y, z) &= f(\rho \cos \theta \sin \phi, \rho \sin \theta \sin \phi, \rho \cos \phi) = \frac{r^3 \cos \theta \sin \theta \sin^2 \phi \cos \phi}{r^3} \\ &= \cos \theta \sin \theta \sin^2 \phi \cos \phi \rightarrow \cos \theta \sin \theta \sin^2 \phi \cos \phi. \end{aligned}$$

Άρα η τιμή του ορίου εξαρτάται από τις γωνίες θ, ϕ και έχει διαφορετικές τιμές όταν αλλάζουμε την ημιευθεία προσέγγισης του (x, y, z) στο $(0, 0, 0)$.

9. Το

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 y}{x^4 + y^2}$$

δεν υπάρχει. Δοκιμάζουμε τις τιμές της συνάρτησης $f(x, y) = \frac{x^2 y}{x^4 + y^2}$ όταν το (x, y) πλησιάζει στο $(0, 0)$ κινούμενο πάνω σε ημιευθεία η οποία σχηματίζει πολική γωνία θ με τον θετικό x -άξονα. Χρησιμοποιούμε, λοιπόν, πολικές συντεταγμένες: όταν το (x, y) πλησιάζει το $(0, 0)$ κινούμενο πάνω σε ημιευθεία η οποία σχηματίζει πολική γωνία θ με τον θετικό x -άξονα τότε έχουμε ότι

$$x = r \cos \theta, \quad y = r \sin \theta$$

και $r \rightarrow 0+$ ενώ το θ είναι συγκεκριμένο και σταθερό. Τότε

$$f(x, y) = f(r \cos \theta, r \sin \theta) = \frac{r^3 \cos^2 \theta \sin \theta}{r^4 \cos^4 \theta + r^2 \sin^2 \theta} = \frac{r \cos^2 \theta \sin \theta}{r^2 \cos^4 \theta + \sin^2 \theta}.$$

Επομένως, αν $\sin \theta \neq 0$ (δηλαδή $\theta \neq 0, \pi, 2\pi$), τότε

$$f(x, y) \rightarrow 0.$$

Αν $\sin \theta = 0$, τότε

$$f(x, y) = f(r \cos \theta, r \sin \theta) = \frac{r^3 \cos^2 \theta \sin \theta}{r^4 \cos^4 \theta + r^2 \sin^2 \theta} = 0 \rightarrow 0.$$

Άρα η τιμή του ορίου είναι ίδια και ίση με 0 ανεξάρτητα από την ημιευθεία προσέγγισης του (x, y) στο $(0, 0)$.

Γιατί, λοιπόν, είπαμε ότι το όριο της συνάρτησης δεν υπάρχει;

Αν το (x, y) πλησιάζει το $(0, 0)$ κινούμενο πάνω στην παραβολή με καρτεσιανή εξίσωση $y = x^2$, τότε το σημείο έχει τη μορφή $(x, y) = (x, x^2)$ και $x \rightarrow 0$. Άρα οι αντίστοιχες τιμές της f είναι

$$f(x, x^2) = \frac{x^2 x^2}{x^4 + (x^2)^2} = \frac{1}{2} \rightarrow \frac{1}{2}$$

και αυτό το όριο είναι διαφορετικό από το 0 το οποίο είναι το όριο όταν το (x, y) πλησιάζει το $(0, 0)$ κινούμενο πάνω σε οποιαδήποτε ημιευθεία με κορυφή το $(0, 0)$.

Ορίσαμε τις μερικές παραγώγους μιας πραγματικής συνάρτησης πολλών μεταβλητών (δηλαδή μίας διανυσματικής μεταβλητής). Έστω

$$f : A \rightarrow \mathbb{R}, \quad A \subseteq \mathbb{R}^n \quad (n \geq 2)$$

και έστω $\mathbf{x} = (x_1, \dots, x_n)$ ένα εσωτερικό σημείο του A . Ορίζουμε την **μερική παράγωγο** της f ως προς την j -οστή μεταβλητή στο σημείο x να είναι το όριο (αν υπάρχει)

$$\begin{aligned}\frac{\partial f}{\partial x_j}(\mathbf{x}) &= \frac{\partial f}{\partial x_j}(x_1, \dots, x_n) \\ &= \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{j-1}, x_j + h, x_{j+1}, \dots, x_n) - f(x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n)}{h}.\end{aligned}$$

Δηλαδή κρατάμε σταθερές όλες τις συντεταγμένες $x_1, \dots, x_n$ του $\mathbf{x}$ εκτός από την j -οστή συντεταγμένη και παραγωγίζουμε ως προς αυτήν την συντεταγμένη. Δηλαδή θεωρούμε τη συνάρτηση $f(x_1, \dots, x_{j-1}, x_j, x_{j+1}, \dots, x_n)$ ως συνάρτηση **μίας** μόνο μεταβλητής, της μεταβλητής x_j , και την παραγωγίζουμε όπως ξέρουμε από τον Απειροστικό Λογισμό μίας μεταβλητής.

Μερικά παραδείγματα.

1. Αν $f(x, y) = x^2y + 4xy^2$, τότε

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= 2xy + 4y^2, \\ \frac{\partial f}{\partial y}(x, y) &= x^2 + 8xy.\end{aligned}$$

Αν θέλουμε να εφαρμόσουμε τον ορισμό, γράφουμε

$$\begin{aligned}\frac{f(x+h, y) - f(x, y)}{h} &= \frac{((x+h)^2y + 4(x+h)y^2) - (x^2y + 4xy^2)}{h} \\ &= \frac{(x^2y + 2xhy + h^2y + 4xy^2 + 4hy^2) - (x^2y + 4xy^2)}{h} \\ &= \frac{2xyh + yh^2 + 4y^2h}{h} = 2xy + 4y^2 + yh\end{aligned}$$

και άρα

$$\frac{\partial f}{\partial x}(x, y) = \lim_{h \rightarrow 0} \frac{f(x+h, y) - f(x, y)}{h} = \lim_{h \rightarrow 0} (2xy + 4y^2 + yh) = 2xy + 4y^2.$$

Ομοίως για την άλλη μεταβλητή:

$$\begin{aligned}\frac{f(x, y+h) - f(x, y)}{h} &= \frac{(x^2(y+h) + 4x(y+h)^2) - (x^2y + 4xy^2)}{h} \\ &= \frac{(x^2y + x^2h + 4xy^2 + 8xyh + 4xh^2) - (x^2y + 4xy^2)}{h} \\ &= \frac{x^2h + 8xyh + 4xh^2}{h} = x^2 + 8xy + 4xh\end{aligned}$$

και άρα

$$\frac{\partial f}{\partial y}(x, y) = \lim_{h \rightarrow 0} \frac{f(x, y+h) - f(x, y)}{h} = \lim_{h \rightarrow 0} (x^2 + 8xy + 4xh) = x^2 + 8xy.$$

2. Αν $f(x, y, z) = xye^z + x \sin(x + y^2 + 3xz)$, τότε

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y, z) &= ye^z + \sin(x + y^2 + 3xz) + x(1 + 3z) \cos(x + y^2 + 3xz), \\ \frac{\partial f}{\partial y}(x, y, z) &= xe^z + 2xy \cos(x + y^2 + 3xz), \\ \frac{\partial f}{\partial z}(x, y, z) &= xye^z + 3x^2 \cos(x + y^2 + 3xz).\end{aligned}$$

6η εβδομάδα.

Γυρίσαμε πίσω στην περίπτωση πραγματικής συνάρτησης μίας μεταβλητής $f : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}$, και ξαναείδαμε τις έννοιες της παραγωγισιμότητας της f σε κάποιο εσωτερικό σημείο x_0 του A και της εφαπτόμενης ευθείας στο γράφημα/καμπύλη της f στο σημείο $(x_0, f(x_0))$.

Θυμηθήκαμε ότι για να πούμε ότι η (μη-κατακόρυφη) ευθεία l με καρτεσιανή εξίσωση

$$y = ax + b$$

είναι εφαπτόμενη στο γράφημα της f στο σημείο $(x_0, f(x_0))$ πρέπει να ισχύουν δύο πράγματα. Πρώτον, η ευθεία l πρέπει να περιέχει το σημείο $(x_0, f(x_0))$. Δηλαδή:

$$f(x_0) = ax_0 + b. \quad (1)$$

Δεύτερον, όταν το $x \in A$ πλησιάζει το x_0 πρέπει η κλίση της χορδής του γραφήματος η οποία βρίσκεται ακριβώς πάνω από το διάστημα του x -άξονα ανάμεσα στα σημεία x_0 και x και η κλίση του ευθ. τμήματος πάνω στην εφαπτόμενη ευθεία το οποίο βρίσκεται ακριβώς πάνω από το ίδιο διάστημα του x -άξονα να προσεγγίζουν η μία την άλλη. Η εν λόγω χορδή του γραφήματος της f συνδέει τα σημεία $(x_0, f(x_0))$ και $(x, f(x))$ και το εν λόγω ευθ.τμήμα της εφαπτόμενης ευθείας συνδέει τα σημεία $(x_0, ax_0 + b)$ και $(x, ax + b)$. Άρα οι αντίστοιχες κλίσεις εκφράζονται από τους λόγους

$$\frac{f(x) - f(x_0)}{x - x_0}, \quad \frac{(ax + b) - (ax_0 + b)}{x - x_0}$$

και επομένως πρέπει να ισχύει

$$\lim_{x \rightarrow x_0} \left| \frac{f(x) - f(x_0)}{x - x_0} - \frac{(ax + b) - (ax_0 + b)}{x - x_0} \right| = 0.$$

Λόγω της (1) έχουμε την ισοδύναμη σχέση

$$\lim_{x \rightarrow x_0} \frac{|f(x) - (ax + b)|}{|x - x_0|} = 0. \quad (2)$$

Βλέπουμε, λοιπόν, ότι η προσέγγιση των δύο κλίσεων εκφράζεται από το όριο (2) το οποίο μπορεί να διατυπωθεί ως εξής: ο λόγος της κατακόρυφης απόστασης των σημείων $(x, f(x))$ και $(x, ax + b)$, τα οποία είναι ακριβώς πάνω από το ίδιο σημείο x και το πρώτο στο γράφημα της f και το δεύτερο στην εφαπτόμενη ευθεία, προς την οριζόντια απόσταση των σημείων x και x_0 του x -άξονα πρέπει να τείνει στο 0 όταν το $x \in A$ πλησιάζει το x_0 .

Τώρα η σχέση (2) γράφεται, πάλι λόγω της (1), ισοδύναμα:

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = a. \quad (3)$$

Άρα για να υπάρχει εφαπτόμενη ευθεία l πρέπει να υπάρχει το όριο (3) και να είναι αριθμός: από το όριο αυτό καθορίζεται ο συντελεστής a και από την σχέση (1) καθορίζεται ο συντελεστής b .

Θυμηθήκαμε ότι η f ονομάζεται παραγωγίσιμη στο x_0 ακριβώς όταν υπάρχει το όριο (3) και είναι αριθμός και ότι αυτό το όριο (όταν υπάρχει) ονομάζεται παράγωγος της f στο x_0 και συμβολίζεται $f'(x_0)$:

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}. \quad (4)$$

Συμπέρασμα: υπάρχει (μη-κατακόρυφη) εφαπτόμενη ευθεία στο γράφημα της f στο σημείο $(x_0, f(x_0))$ αν και μόνο αν η f είναι παραγωγίσιμη στο x_0 και τότε οι συντελεστές της καρτεσιανής εξίσωσης της εφαπτόμενης ευθείας είναι οι

$$\begin{aligned} a &= f'(x_0) \\ b &= f(x_0) - ax_0 = f(x_0) - f'(x_0)x_0 \end{aligned}$$

οπότε η καρτεσιανή εξίσωση της εφαπτόμενης ευθείας είναι η

$$y = f'(x_0)(x - x_0) + f(x_0).$$

Το όριο (4) το οποίο καθορίζει την παραγωγισιμότητα της f στο x_0 γράφεται ισοδύναμα

$$\lim_{x \rightarrow x_0} \left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| = 0$$

ή, ισοδύναμα,

$$\lim_{x \rightarrow x_0} \frac{|f(x) - f(x_0) - f'(x_0)(x - x_0)|}{|x - x_0|} = 0. \quad (5)$$

Παρατηρήστε ότι η σχέσεις (2) και (5) είναι ταυτόσημες λόγω των συγκεκριμένων τιμών των συντελεστών a, b .

Έχοντας θυμηθεί την περίπτωση πραγματικής συνάρτησης μίας μεταβλητής, πάμε στην περίπτωση πραγματικής συνάρτησης δύο (πραγματικών) μεταβλητών.

Θεωρούμε $f : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^2$ και εσωτερικό σημείο (x_0, y_0) του A . Τώρα το γράφημα μιας τέτοιας συνάρτησης είναι συνήθως μια επιφάνεια στον $\mathbb{R}^3$ και θα ασχοληθούμε με την ύπαρξη εφαπτόμενου επιπέδου στο γράφημα/επιφάνεια της f στο σημείο $(x_0, y_0, f(x_0, y_0))$. Για να πούμε ότι το (μη-κατακόρυφο) επίπεδο l με καρτεσιανή εξίσωση

$$z = ax + by + c$$

είναι εφαπτόμενο στο γράφημα της f στο σημείο $(x_0, y_0, f(x_0, y_0))$ πρέπει να ισχύουν δύο πράγματα. Πρώτον, το επίπεδο l πρέπει να περιέχει το σημείο $(x_0, y_0, f(x_0, y_0))$. Δηλαδή:

$$f(x_0) = ax_0 + by_0 + c. \quad (6)$$

Δεύτερον, όταν το $(x, y) \in A$ πλησιάζει το (x_0, y_0) πρέπει η κλίση της χορδής του γραφήματος η οποία βρίσκεται ακριβώς πάνω από το ευθ. τμήμα του xy -επιπέδου ανάμεσα στα σημεία (x_0, y_0) και (x, y) και η κλίση του ευθ. τμήματος πάνω στο εφαπτόμενο επίπεδο το οποίο βρίσκεται ακριβώς πάνω από το ίδιο ευθ. τμήμα του xy -επιπέδου να προσεγγίζουν η μία την άλλη. Η εν λόγω χορδή του γραφήματος της f συνδέει τα σημεία $(x_0, y_0, f(x_0, y_0))$ και $(x, y, f(x, y))$ και το εν λόγω ευθ.τμήμα του εφαπτόμενου επιπέδου συνδέει τα σημεία $(x_0, y_0, ax_0 + by_0 + c)$ και $(x, y, ax + by + c)$. Η προσέγγιση των δύο κλίσεων (όπως και στην περίπτωση της μίας μεταβλητής) μπορεί να διατυπωθεί ως εξής: ο λόγος της κατακόρυφης απόστασης των σημείων $(x, y, f(x, y))$ και $(x, y, ax + by + c)$, τα οποία είναι ακριβώς πάνω από το ίδιο σημείο (x, y) και το πρώτο στο γράφημα της f και το δεύτερο στο εφαπτόμενο επίπεδο, προς την οριζόντια απόσταση των σημείων (x, y) και (x_0, y_0) του xy -επιπέδου πρέπει να τείνει στο 0 όταν το $(x, y) \in A$ πλησιάζει το (x_0, y_0) . Άρα η προσέγγιση των δύο κλίσεων ισοδυναμεί με το όριο

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{|f(x, y) - (ax + by + c)|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0. \quad (7)$$

Οι σχέσεις (6) και (7) είναι ανάλογες των (1) και (2).

Χρησιμοποιώντας την (6), αντικαθιστούμε το c στην (7) και έχουμε την ισοδύναμη

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{|f(x,y) - f(x_0,y_0) - a(x-x_0) - b(y-y_0)|}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} = 0. \quad (8)$$

Βλέπουμε, λοιπόν, ότι το να υπάρχει εφαπτόμενο επίπεδο στο γράφημα της f στο σημείο $(x_0, y_0, f(x_0, y_0))$ ισοδυναμεί με τις σχέσεις (6) και (8). Από την (8) προσδιορίζονται (αν υπάρχουν) οι συντελεστές a, b και κατόπιν από την (6) ο συντελεστής c της καρτεσιανής εξίσωσης του εφαπτόμενου επιπέδου.

Από την (8) συνεπάγονται (αλλά όχι αντίστροφα) δύο αναγκαίες σχέσεις ως εξής. Αν το σημείο (x, y) πλησιάζει το (x_0, y_0) κινούμενο παράλληλα στον x -άξονα, τότε έχουμε $(x, y) = (x, y_0)$ και $x \rightarrow x_0$ και από την (8) παίρνουμε ότι

$$\lim_{x \rightarrow x_0} \frac{|f(x, y_0) - f(x_0, y_0) - a(x - x_0)|}{|x - x_0|} = 0$$

ή, ισοδύναμα:

$$\lim_{x \rightarrow x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} = a. \quad (9)$$

Ομοίως, αν το σημείο (x, y) πλησιάζει το (x_0, y_0) κινούμενο παράλληλα στον y -άξονα, τότε έχουμε $(x, y) = (x_0, y)$ και $y \rightarrow y_0$ και από την (8) παίρνουμε ότι

$$\lim_{y \rightarrow y_0} \frac{|f(x_0, y) - f(x_0, y_0) - b(y - y_0)|}{|y - y_0|} = 0$$

ή, ισοδύναμα:

$$\lim_{y \rightarrow y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0} = b. \quad (10)$$

Σύμφωνα με τον ορισμό των μερικών παραγώγων, οι (9), (10) σημαίνουν ότι

$$\frac{\partial f}{\partial x}(x_0, y_0) = a, \quad \frac{\partial f}{\partial y}(x_0, y_0) = b.$$

Άρα, αν υπάρχει το εφαπτόμενο επίπεδο l , τότε συνεπάγεται ότι υπάρχουν οι μερικές παράγωγοι της f ως προς x και ως προς y στο (x_0, y_0) και οι συντελεστές a, b είναι ίσοι με αυτές τις μερικές παραγώγους. Από την σχέση (6) καθορίζεται και το c και άρα οι συντελεστές της καρτεσιανής εξίσωσης του εφαπτόμενου επιπέδου είναι οι

$$a = \frac{\partial f}{\partial x}(x_0, y_0)$$

$$b = \frac{\partial f}{\partial y}(x_0, y_0)$$

$$c = f(x_0, y_0) - ax_0 - by_0 = f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)x_0 - \frac{\partial f}{\partial y}(x_0, y_0)y_0$$

και η καρτεσιανή εξίσωσή του είναι η

$$z = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + f(x_0, y_0). \quad (11)$$

Μέχρι στιγμής έχουμε συμπεράνει τα εξής. Οι σχέσεις (6) και (8) είναι ισοδύναμες με την ύπαρξη του εφαπτόμενου επιπέδου και από την (8) συνεπάγεται η ύπαρξη των μερικών παραγώγων της f στο (x_0, y_0) . Η ύπαρξη των μερικών παραγώγων προέκυψε από την (8) παίρνοντας όριο όταν το (x, y) πλησιάζει το (x_0, y_0) πάνω σε δύο μόνο “δρόμους”: έναν παράλληλο στον

x -άξονα και έναν παράλληλο στον y -άξονα. Όμως, γνωρίζουμε ήδη ότι το όριο (8) δεν ισοδυναμεί με το όριο από δύο μόνο “δρόμους”. Επομένως, η ύπαρξη των μερικών παραγώγων στο (x_0, y_0) δεν συνεπάγεται το όριο (8). Βάσει, όμως, του ορίου (8) έχουμε καταφέρει να βρούμε τους συντελεστές και την εξίσωση του υποψήφιου εφαπτόμενου επιπέδου και άρα το όριο (8) μπορεί να επαναδιατυπωθεί στην πιο συγκεκριμένη μορφή:

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{|f(x,y) - f(x_0,y_0) - [\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)]|}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} = 0. \quad (12)$$

Επαναδιατυπώνουμε και τα μέχρι στιγμής συμπεράσματά μας. Υπάρχει το (μη-κατακόρυφο) εφαπτόμενο επίπεδο στο γράφημα της f στο σημείο $(x_0, y_0, f(x_0, y_0))$ αν και μόνο αν υπάρχουν οι μερικές παράγωγοι της f ως προς x και ως προς y στο (x_0, y_0) και ισχύει το όριο (12). Σ’ αυτήν την περίπτωση η καρτεσιανή εξίσωση του εφαπτόμενου επιπέδου είναι η (11).

Τώρα δίνουμε το εξής σημαντικό ορισμό.

Λέμε ότι η f είναι **παραγωγίσιμη** στο (x_0, y_0) όταν υπάρχουν οι μερικές παράγωγοι της f ως προς x και ως προς y στο (x_0, y_0) και ισχύει το όριο (12).

Βάσει του τελευταίου συμπεράσματος και του ορισμού της παραγωγισιμότητας που δώσαμε, το επόμενο συμπέρασμα είναι απλώς μια ταυτολογία: υπάρχει το (μη-κατακόρυφο) εφαπτόμενο επίπεδο στο γράφημα της f στο σημείο $(x_0, y_0, f(x_0, y_0))$ αν και μόνο αν η f είναι παραγωγίσιμη στο (x_0, y_0) .

Δύο παραδείγματα:

1. Θεωρούμε την συνάρτηση

$$f(x, y) = xy.$$

Θέλουμε να δούμε αν υπάρχει εφαπτόμενο επίπεδο στο γράφημα της f στο σημείο $(1, 1, f(1, 1)) = (1, 1, 1)$ ή, ισοδύναμα, αν η f είναι παραγωγίσιμη στο $(1, 1)$. Βρίσκουμε εύκολα τις μερικές παραγώγους

$$\frac{\partial f}{\partial x}(x, y) = y, \quad \frac{\partial f}{\partial y}(x, y) = x$$

και άρα

$$\frac{\partial f}{\partial x}(1, 1) = 1, \quad \frac{\partial f}{\partial y}(1, 1) = 1.$$

Άρα πρέπει να ελέγξουμε το όριο

$$\lim_{(x,y) \rightarrow (1,1)} \frac{|f(x,y) - f(1,1) - [\frac{\partial f}{\partial x}(1,1)(x-1) + \frac{\partial f}{\partial y}(1,1)(y-1)]|}{\sqrt{(x-1)^2 + (y-1)^2}} = 0$$

δηλαδή το

$$\lim_{(x,y) \rightarrow (1,1)} \frac{|xy - 1 - [(x-1) + (y-1)]|}{\sqrt{(x-1)^2 + (y-1)^2}} = 0$$

δηλαδή, μετά από λίγες πράξεις και μια απλή παραγοντοποίηση, το

$$\lim_{(x,y) \rightarrow (1,1)} \frac{|(x-1)(y-1)|}{\sqrt{(x-1)^2 + (y-1)^2}} = 0.$$

Κάνουμε μια απλή και προφανή αλλαγή μεταβλητής:

$$u = x - 1, \quad v = y - 1$$

και το όριο που θέλουμε να αποδείξουμε γράφεται

$$\lim_{(u,v) \rightarrow (0,0)} \frac{|uv|}{\sqrt{u^2 + v^2}} = 0.$$

Χρησιμοποιώντας την απλή ανισότητα

$$0 \leq \frac{|uv|}{\sqrt{u^2 + v^2}} \leq |v|$$

βλέπουμε ότι το όριο είναι σωστό και άρα η f είναι παραγωγίσιμη στο $(1, 1)$. Η καρτεσιανή εξίσωση του εφαπτόμενου επιπέδου είναι

$$z = \frac{\partial f}{\partial x}(1, 1)(x - 1) + \frac{\partial f}{\partial y}(1, 1)(y - 1) + f(1, 1) = (x - 1) + (y - 1) + 1 = x + y - 1.$$

2. Θεωρούμε την συνάρτηση

$$f(x, y) = \sqrt{|xy|}$$

και θέλουμε πάλι να δούμε αν υπάρχει το εφαπτόμενο επίπεδο στο γράφημα της f στο σημείο $(0, 0, f(0, 0)) = (0, 0, 0)$ ή, ισοδύναμα, αν η f είναι παραγωγίσιμη στο $(0, 0)$. Βρίσκουμε τις μερικές παραγώγους

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = \lim_{h \rightarrow 0} 0 = 0,$$

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = \lim_{h \rightarrow 0} 0 = 0.$$

Άρα πρέπει να ελέγξουμε το όριο

$$\lim_{(x, y) \rightarrow (0, 0)} \frac{|f(x, y) - f(0, 0) - [\frac{\partial f}{\partial x}(0, 0)(x - 0) + \frac{\partial f}{\partial y}(0, 0)(y - 0)]|}{\sqrt{(x - 0)^2 + (y - 0)^2}} = 0$$

δηλαδή το

$$\lim_{(x, y) \rightarrow (0, 0)} \frac{\sqrt{|xy|}}{\sqrt{x^2 + y^2}} = 0$$

δηλαδή το

$$\lim_{(x, y) \rightarrow (0, 0)} \frac{|xy|}{x^2 + y^2} = 0.$$

Βλέπουμε, όμως, ότι όταν το (x, y) πλησιάζει το $(0, 0)$ κινούμενο πάνω στην κύρια διαγώνιο με καρτεσιανή εξίσωση $y = x$ έχουμε ότι $(x, y) = (x, x)$ και $x \rightarrow 0$ και τότε

$$\frac{|xy|}{x^2 + y^2} = \frac{x^2}{2x^2} = \frac{1}{2} \rightarrow \frac{1}{2} \neq 0.$$

Άρα η f δεν είναι παραγωγίσιμη στο $(0, 0)$.

Υπάρχει το εξής **Κριτήριο Παραγωγισιμότητας**.

Αν η f έχει μερικές παραγώγους σε κάθε σημείο (x, y) ενός δίσκου με κέντρο το (x_0, y_0) και θετική ακτίνα και αν αυτές οι μερικές παράγωγοι είναι συνεχείς στο (x_0, y_0) , τότε η f είναι παραγωγίσιμη στο (x_0, y_0) .

Το κριτήριο παραγωγισιμότητας μας επιτρέπει να βλέπουμε εύκολα αν μια συνάρτηση είναι παραγωγίσιμη. Ας ξαναδούμε το πρώτο από τα δύο προηγούμενα παραδείγματα.

3. Θεωρούμε την συνάρτηση

$$f(x, y) = xy.$$

Η f έχει σε κάθε σημείο (x, y) του xy -επιπέδου μερικές παραγώγους

$$\frac{\partial f}{\partial x}(x, y) = y, \quad \frac{\partial f}{\partial y}(x, y) = x$$

οι οποίες είναι, προφανώς, συνεχείς σε κάθε σημείο (x_0, y_0) του xy -επιπέδου. Άρα η f είναι παραγωγίσιμη σε κάθε σημείο (x_0, y_0) και όχι μόνο στο $(1, 1)$ που είδαμε προηγουμένως. Η εξίσωση του εφαπτόμενου επιπέδου στο γράφημα της f στο τυχόν σημείο $(x_0, y_0, f(x_0, y_0)) = (x_0, y_0, x_0 y_0)$ είναι η

$$\begin{aligned} z &= \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + f(x_0, y_0) = y_0(x - x_0) + x_0(y - y_0) + x_0 y_0 \\ &= y_0 x + x_0 y - x_0 y_0. \end{aligned}$$

Δίνουμε τον ορισμό της παραγωγισιμότητας για πραγματικές συναρτήσεις πολλών (πραγματικών) μεταβλητών.

Έστω $f : A \rightarrow \mathbb{R}$ όπου $A \subseteq \mathbb{R}^n$ ($n \geq 2$) και έστω $\mathbf{x}_0$ εσωτερικό σημείο του A . Λέμε ότι η f είναι **παραγωγίσιμη** στο $\mathbf{x}_0$ όταν υπάρχουν οι μερικές παράγωγοι της f ως προς $x_1, \dots, x_n$ στο $\mathbf{x}_0$ και ισχύει το όριο

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{|f(\mathbf{x}) - f(\mathbf{x}_0) - [\frac{\partial f}{\partial x_1}(\mathbf{x}_0)(x_1 - x_{01}) + \dots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0)(x_n - x_{0n})]|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0. \quad (13)$$

($\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{x}_0 = (x_{01}, \dots, x_{0n})$, $\|\mathbf{x} - \mathbf{x}_0\| = \sqrt{(x_1 - x_{01})^2 + \dots + (x_n - x_{0n})^2}$.)

Υπάρχει και το ανάλογο **Κριτήριο Παραγωγισιμότητας**.

Αν η f έχει μερικές παραγώγους σε κάθε σημείο $\mathbf{x}$ μιας μπάλας με κέντρο το $\mathbf{x}_0$ και θετική ακτίνα και αν αυτές οι μερικές παράγωγοι είναι συνεχείς στο $\mathbf{x}_0$, τότε η f είναι παραγωγίσιμη στο $\mathbf{x}_0$.

Παράδειγμα.

4. Θεωρούμε την συνάρτηση

$$f(x, y, z) = x e^{yz}.$$

Η f έχει σε κάθε σημείο (x, y, z) του $\mathbb{R}^3$ μερικές παραγώγους

$$\frac{\partial f}{\partial x}(x, y, z) = e^{yz}, \quad \frac{\partial f}{\partial y}(x, y, z) = x z e^{yz}, \quad \frac{\partial f}{\partial z}(x, y, z) = x y e^{yz}$$

οι οποίες είναι συνεχείς σε κάθε σημείο (x_0, y_0, z_0) του $\mathbb{R}^3$. Άρα η f είναι παραγωγίσιμη σε κάθε σημείο (x_0, y_0, z_0) . Η εξίσωση του εφαπτόμενου “επιπέδου” στο γράφημα της f στο τυχόν σημείο $(x_0, y_0, z_0, f(x_0, y_0, z_0)) = (x_0, y_0, z_0, x_0 e^{y_0 z_0})$ είναι η

$$\begin{aligned} w &= \frac{\partial f}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0, z_0)(y - y_0) + \frac{\partial f}{\partial z}(x_0, y_0, z_0)(z - z_0) + f(x_0, y_0, z_0) \\ &= e^{y_0 z_0}(x - x_0) + x_0 z_0 e^{y_0 z_0}(y - y_0) + x_0 y_0 e^{y_0 z_0}(z - z_0) + x_0 e^{y_0 z_0}. \end{aligned}$$

Λέγοντας εφαπτόμενο “επίπεδο” βάζουμε εισαγωγικά διότι το γράφημα της f βρίσκεται στον $\mathbb{R}^4$ και είναι μία τρισδιάστατη επιφάνεια (έναν τρισδιάστατο χώρο μέσα στον τετραδιάστατο χώρο) και άρα στα σημεία του έχει τρισδιάστατα εφαπτόμενα “επίπεδα” (δηλαδή τρισδιάστατους εφαπτόμενους χώρους).

Έστω $f : A \rightarrow \mathbb{R}$ όπου $A \subseteq \mathbb{R}^n$ ($n \geq 2$) και έστω $\mathbf{x}_0$ εσωτερικό σημείο του A . Αν υπάρχουν οι μερικές παράγωγοι της f ως προς $x_1, \dots, x_n$ στο $\mathbf{x}_0$ συμβολίζουμε

$$\begin{aligned} \nabla f(\mathbf{x}_0) &= \text{grad } f(\mathbf{x}_0) = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}_0), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right), \\ Df(\mathbf{x}_0) &= \left[\frac{\partial f}{\partial x_1}(\mathbf{x}_0) \cdots \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \right]. \end{aligned}$$

Το $\nabla f(\mathbf{x}_0)$ ή $\text{grad } f(\mathbf{x}_0)$ ονομάζεται **κλίση** της f στο $\mathbf{x}_0$ και είναι διάνυσμα στον $\mathbb{R}^n$. Το $Df(\mathbf{x}_0)$ ονομάζεται **παράγωγος** της f στο $\mathbf{x}_0$ και είναι $1 \times n$ πίνακας.

Έτσι η παράσταση

$$\frac{\partial f}{\partial x_1}(\mathbf{x}_0)(x_1 - x_{01}) + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0)(x_n - x_{0n})$$

που εμφανίζεται στο όριο (13) γράφεται με δύο τρόπους ως εξής. Ως εσωτερικό γινόμενο διανυσμάτων στον $\mathbb{R}^n$:

$$\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

Και ως γινόμενο ενός $1 \times n$ πίνακα με έναν $n \times 1$ πίνακα:

$$Df(\mathbf{x}_0) \begin{bmatrix} x_1 - x_{01} \\ \vdots \\ x_n - x_{0n} \end{bmatrix} = Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

Στην πρώτη παράσταση το $\mathbf{x} - \mathbf{x}_0$ είναι διάνυσμα στον $\mathbb{R}^n$ ενώ στην δεύτερη παράσταση το $\mathbf{x} - \mathbf{x}_0$ είναι πίνακας-στήλη. Επίσης πρέπει να πούμε ότι στην δεύτερη παράσταση το αποτέλεσμα του πολλαπλασιασμού των πινάκων είναι ένας 1×1 πίνακας και ότι θεωρούμε έναν 1×1 πίνακα ως αριθμό.

Άρα το όριο (13) γράφεται και ως εξής:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{|f(\mathbf{x}) - f(\mathbf{x}_0) - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$$

και

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{|f(\mathbf{x}) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

7η εβδομάδα.

Γενικεύσαμε τη συζήτηση περί παραγωγισιμότητας από πραγματικές σε διανυσματικές συναρτήσεις:

$$\mathbf{f} : A \rightarrow \mathbb{R}^m, \quad A \subseteq \mathbb{R}^n, \quad (14)$$

όπου $m \geq 2$.

Προηγουμένως μιλήσαμε για τις γραμμικές συναρτήσεις που μαθαίνουμε στην Γραμμική Άλγεβρα. Αν

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$$

είναι ένας $m \times n$ πίνακας, τότε ορίζεται η αντίστοιχη γραμμική συνάρτηση

$$\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

με τύπο

$$\mathbf{y} = \mathbf{f}(\mathbf{x}) = A\mathbf{x}, \quad (15)$$

όπου η δεξιά μεριά είναι το γινόμενο του $m \times n$ πίνακα A και του $n \times 1$ πίνακα $\mathbf{x}$ και άρα το $\mathbf{y} = \mathbf{f}(\mathbf{x})$ είναι $m \times 1$ πίνακας. Τότε

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

και άρα οι συντεταγμένες συναρτήσεις της $\mathbf{f}$ είναι οι

$$\begin{aligned} y_1 &= f_1(x_1, \dots, x_n) = a_{11}x_1 + \dots + a_{1n}x_n \\ &\vdots \\ y_m &= f_m(x_1, \dots, x_n) = a_{m1}x_1 + \dots + a_{mn}x_n \end{aligned} \quad (16)$$

Πηγαίνοντας στην γενική περίπτωση διανυσματικής συνάρτησης (14), θεωρούμε τις συντεταγμένες συναρτήσεις της:

$$f_i : A \rightarrow \mathbb{R}, \quad 1 \leq i \leq m.$$

Οι $f_1, \dots, f_m$ είναι πραγματικές συναρτήσεις και θεωρούμε ότι σε κάποιο εσωτερικό σημείο $\mathbf{x}_0$ του A έχουν μερικές παραγώγους

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0), \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

Σχηματίζουμε τον $m \times n$ πίνακα

$$D\mathbf{f}(\mathbf{x}_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}_0) \end{bmatrix}$$

ο οποίος στην i -οστή γραμμή έχει τις μερικές παραγώγους της i -οστής συνάρτησης f_i . Αν η $\mathbf{f}$ ήταν πραγματική συνάρτηση, θα είχε μόνο μία συντεταγμένη συνάρτηση, τον εαυτό της, και ο πίνακας $D\mathbf{f}(\mathbf{x}_0)$ θα ήταν $1 \times n$ πίνακας -αυτός που είδαμε την προηγούμενη εβδομάδα.

Τώρα, λέμε ότι η $\mathbf{f}$ είναι **παραγωγίσιμη** στο $\mathbf{x}_0$ αν ισχύει το όριο:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0. \quad (17)$$

Το όριο αυτό μοιάζει με το τελευταίο όριο της προηγούμενης εβδομάδας το οποίο καθορίζει την παραγωγισιμότητα πραγματικής συνάρτησης f :

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{|f(\mathbf{x}) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

Στην περίπτωση πραγματικής συνάρτησης f η παράσταση $f(\mathbf{x}) - f(\mathbf{x}_0) - Df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ είναι πραγματικός αριθμός και γι αυτό μπορούμε να θεωρήσουμε την απόλυτη τιμή της. Στην περίπτωση διανυσματικής συνάρτησης $\mathbf{f}$ η ανάλογη παράσταση $\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)$ (στην (17)) είναι $m \times 1$ πίνακας -δηλαδή στοιχείο του $\mathbb{R}^m$ - και θεωρούμε το μήκος της (γενίκευση της απόλυτης τιμής από την μία στις πολλές διαστάσεις). Επομένως ο ορισμός της παραγωγισιμότητας διανυσματικής συνάρτησης είναι κατ' ευθείαν γενίκευση του ορισμού της παραγωγισιμότητας πραγματικής συνάρτησης.

Αν η $\mathbf{f}$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ (δηλαδή αν ισχύει το όριο (17)) ο πίνακας $D\mathbf{f}(\mathbf{x}_0)$ ονομάζεται **παράγωγος** της $\mathbf{f}$ στο $\mathbf{x}_0$.

Δύο απλά παραδείγματα.

1. Μια σταθερή συνάρτηση $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ με τύπο

$$\mathbf{f}(\mathbf{x}) = \mathbf{y}_0 \quad \text{για κάθε } \mathbf{x} \in \mathbb{R}^n.$$

Κάθε συντεταγμένη συνάρτηση της $\mathbf{f}$ είναι σταθερή:

$$f_i(\mathbf{x}) = y_{0i} \quad \text{για κάθε } \mathbf{x} \in \mathbb{R}^n.$$

Άρα οι μερικές παράγωγοι όλων των f_i είναι 0 σε κάθε σημείο $\mathbf{x}_0$ και άρα ο πίνακας $D\mathbf{f}(\mathbf{x}_0)$ είναι ο μηδενικός $m \times n$ πίνακας:

$$D\mathbf{f}(\mathbf{x}_0) = \mathbf{0} = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{bmatrix}$$

Άρα

$$\begin{aligned} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} &= \frac{\|\mathbf{y}_0 - \mathbf{y}_0 - \mathbf{0}(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = \frac{\|\mathbf{y}_0 - \mathbf{y}_0 - \mathbf{0}\|}{\|\mathbf{x} - \mathbf{x}_0\|} = \frac{\|\mathbf{0}\|}{\|\mathbf{x} - \mathbf{x}_0\|} \\ &= 0 \rightarrow 0 \quad \text{όταν } \mathbf{x} \rightarrow \mathbf{x}_0. \end{aligned}$$

Άρα κάθε σταθερή συνάρτηση $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμη σε κάθε σημείο $\mathbf{x}_0$ και έχει παράγωγο τον μηδενικό $m \times n$ πίνακα $\mathbf{0}$.

2. Θεωρούμε την γραμμική συνάρτηση που ορίζεται από έναν $m \times n$ πίνακα A όπως στην σχέση (15). Από τους τύπους (16) των συντεταγμένων συναρτήσεων βλέπουμε ότι

$$\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) = a_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n,$$

σε κάθε σημείο $\mathbf{x}_0$ και άρα

$$D\mathbf{f}(\mathbf{x}_0) = A.$$

Τότε

$$\begin{aligned}\frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}_0) - D\mathbf{f}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} &= \frac{\|A\mathbf{x} - A\mathbf{x}_0 - A(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = \frac{\|A(\mathbf{x} - \mathbf{x}_0) - A(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} \\ &= \frac{\|\mathbf{0}\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0 \rightarrow 0 \quad \text{όταν } \mathbf{x} \rightarrow \mathbf{x}_0.\end{aligned}$$

(Χρησιμοποιήσαμε την επιμεριστική ιδιότητα του πολλαπλασιασμού πινάκων.) Άρα κάθε γραμμική συνάρτηση $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμη σε κάθε σημείο $\mathbf{x}_0$ και έχει παράγωγο τον $m \times n$ πίνακα A ο οποίος την καθορίζει.

Αναφέραμε διάφορες ιδιότητες της παραγωγισιμότητας. (Προσέξτε τις διαστάσεις των διαφόρων πινάκων.)

1. Αν η $\mathbf{f} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμη στο $\mathbf{x}_0$, τότε είναι συνεχής στο $\mathbf{x}_0$.
2. Αν οι $\mathbf{f}, \mathbf{g} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμες στο $\mathbf{x}_0$, τότε και οι $\mathbf{f} \pm \mathbf{g} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμες στο $\mathbf{x}_0$ και

$$D(\mathbf{f} \pm \mathbf{g})(\mathbf{x}_0) = D\mathbf{f}(\mathbf{x}_0) \pm D\mathbf{g}(\mathbf{x}_0).$$

3. Αν η $\mathbf{f} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ και αν $\lambda \in \mathbb{R}$ τότε και η $\lambda\mathbf{f} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ και

$$D(\lambda\mathbf{f})(\mathbf{x}_0) = \lambda D\mathbf{f}(\mathbf{x}_0).$$

4. Αν οι $f : A \rightarrow \mathbb{R}$ και $\mathbf{g} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμες στο $\mathbf{x}_0$, τότε και η $f\mathbf{g} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ και

$$D(f\mathbf{g})(\mathbf{x}_0) = f(\mathbf{x}_0)D\mathbf{g}(\mathbf{x}_0) + \mathbf{g}(\mathbf{x}_0)Df(\mathbf{x}_0).$$

(Το $\mathbf{g}(\mathbf{x}_0)$ είναι $m \times 1$ πίνακας.)

5. Αν οι $f : A \rightarrow \mathbb{R}$ και $\mathbf{g} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμες στο $\mathbf{x}_0$ και αν $f(\mathbf{x}) \neq 0$ για κάθε $\mathbf{x} \in A$, τότε και η $\frac{\mathbf{g}}{f} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ και

$$D(\mathbf{g}/f)(\mathbf{x}_0) = (f(\mathbf{x}_0)D\mathbf{g}(\mathbf{x}_0) - \mathbf{g}(\mathbf{x}_0)Df(\mathbf{x}_0))/(f(\mathbf{x}_0))^2.$$

(Το $\mathbf{g}(\mathbf{x}_0)$ είναι $m \times 1$ πίνακας.)

6. Αν οι $\mathbf{f}, \mathbf{g} : A \rightarrow \mathbb{R}^m$ είναι παραγωγίσιμες στο $\mathbf{x}_0$, τότε και η $\mathbf{f} \cdot \mathbf{g} : A \rightarrow \mathbb{R}$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ και

$$D(\mathbf{f} \cdot \mathbf{g})(\mathbf{x}_0) = \mathbf{f}(\mathbf{x}_0)D\mathbf{g}(\mathbf{x}_0) + \mathbf{g}(\mathbf{x}_0)D\mathbf{f}(\mathbf{x}_0).$$

(Τα $\mathbf{f}(\mathbf{x}_0), \mathbf{g}(\mathbf{x}_0)$ είναι $1 \times m$ πίνακες.)

7. Έστω $\mathbf{f} : A \rightarrow \mathbb{R}^m$ και έστω $f_1, \dots, f_m : A \rightarrow \mathbb{R}$ οι συντεταγμένες συναρτήσεις της $\mathbf{f}$. Τότε η $\mathbf{f}$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ αν και μόνο αν όλες οι $f_1, \dots, f_m$ είναι παραγωγίσιμες στο $\mathbf{x}_0$.

8. Κριτήριο Παραγωγισιμότητας. Αν καθεμία από τις συντεταγμένες συναρτήσεις $f_1, \dots, f_m$ της $\mathbf{f}$ έχει μερικές παραγώγους σε κάθε σημείο $\mathbf{x}$ μιας μπάλας με κέντρο το $\mathbf{x}_0$ και θετική ακτίνα και αν αυτές οι μερικές παράγωγοι είναι συνεχείς στο $\mathbf{x}_0$, τότε η $\mathbf{f}$ είναι παραγωγίσιμη στο $\mathbf{x}_0$.

Παράδειγμα. Η συνάρτηση $\mathbf{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ με τύπο

$$f(x, y, z) = (x^2y + yz, e^{x+2y+z^2})$$

έχει συντεταγμένες συναρτήσεις $f_1, f_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$ με τύπους

$$f_1(x, y, z) = x^2y + yz, \quad f_2(x, y, z) = e^{x+2y+z^2}.$$

Οι μερικές παράγωγοι των f_1, f_2 είναι οι

$$\frac{\partial f_1}{\partial x}(x, y) = 2xy, \quad \frac{\partial f_1}{\partial y}(x, y) = x^2 + z, \quad \frac{\partial f_1}{\partial z}(x, y) = y,$$

$$\frac{\partial f_2}{\partial x}(x, y) = e^{x+2y+z^2}, \quad \frac{\partial f_2}{\partial y}(x, y) = 2e^{x+2y+z^2}, \quad \frac{\partial f_2}{\partial z}(x, y) = 2ze^{x+2y+z^2}$$

και είναι συνεχείς σε κάθε σημείο του $\mathbb{R}^3$. Άρα η $\mathbf{f}$ είναι παραγωγίσιμη σε κάθε σημείο του $\mathbb{R}^3$.

9. Κανόνας Αλυσίδας. Αν $\mathbf{f} : A \rightarrow B$, όπου $A \subseteq \mathbb{R}^n$ και $B \subseteq \mathbb{R}^m$, και $\mathbf{g} : B \rightarrow \mathbb{R}^k$ και η $\mathbf{f}$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ και η $\mathbf{g}$ είναι παραγωγίσιμη στο $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$ τότε και η $\mathbf{h} = \mathbf{g} \circ \mathbf{f} : A \rightarrow \mathbb{R}^k$ είναι παραγωγίσιμη στο $\mathbf{x}_0$ και

$$D\mathbf{h}(\mathbf{x}_0) = D\mathbf{g}(\mathbf{y}_0)D\mathbf{f}(\mathbf{x}_0), \quad \mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0). \quad (18)$$

Ο κανόνας της αλυσίδας είναι πολύ σημαντικός και πρέπει να δοθεί ιδιαίτερη προσοχή σ' αυτόν.

Κατ' αρχάς παρατηρούμε την ομοιότητα με τον κανόνα αλυσίδας για πραγματικές συναρτήσεις μίας μεταβλητής:

$$h'(x_0) = g'(y_0)f'(x_0), \quad y_0 = f(x_0).$$

Είναι σαφές ότι η σχέση (18) αποτελεί άμεση γενίκευση της προηγούμενης σχέσης.

Πρέπει να κατανοήσουμε τον χαρακτήρα της σχέσης (18). Η αριστερή μεριά είναι ένας $k \times n$ πίνακας. Η δεξιά μεριά είναι γινόμενο ενός $k \times m$ πίνακα και ενός $m \times n$ πίνακα δηλαδή ένας $k \times n$ πίνακας. Άρα η σχέση (18) κατ' αρχάς έχει νόημα.

Για να κατανοήσουμε καλύτερα την ισότητα (18) θα δούμε ποιά είναι η (l, j) -οστή συντεταγμένη του πίνακα στην αριστερή μεριά και ποιά είναι η (l, j) -οστή συντεταγμένη του πίνακα που προκύπτει μετά από τον πολλαπλασιασμό των πινάκων στην δεξιά μεριά. Η (l, j) -οστή συντεταγμένη του πίνακα στην αριστερή μεριά είναι η μερική παράγωγος της συντεταγμένης συνάρτησης h_l ως προς την μεταβλητή x_j . Δηλαδή

$$\frac{\partial h_l}{\partial x_j}(\mathbf{x}_0).$$

Η (l, j) -οστή συντεταγμένη του πίνακα που προκύπτει μετά από τον πολλαπλασιασμό των πινάκων στην δεξιά μεριά είναι το εσωτερικό γινόμενο της l -οστής γραμμής του πίνακα $D\mathbf{g}(\mathbf{y}_0)$ και της j -οστής στήλης του πίνακα $D\mathbf{f}(\mathbf{x}_0)$. Δηλαδή

$$\frac{\partial g_l}{\partial y_1}(\mathbf{y}_0) \frac{\partial f_1}{\partial x_j}(\mathbf{x}_0) + \cdots + \frac{\partial g_l}{\partial y_m}(\mathbf{y}_0) \frac{\partial f_m}{\partial x_j}(\mathbf{x}_0), \quad \mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0).$$

Άρα η ισότητα (18) ισοδυναμεί με

$$\frac{\partial h_l}{\partial x_j}(\mathbf{x}_0) = \frac{\partial g_l}{\partial y_1}(\mathbf{y}_0) \frac{\partial f_1}{\partial x_j}(\mathbf{x}_0) + \cdots + \frac{\partial g_l}{\partial y_m}(\mathbf{y}_0) \frac{\partial f_m}{\partial x_j}(\mathbf{x}_0), \quad 1 \leq l \leq k, \quad 1 \leq j \leq n, \quad \mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0) \quad (19)$$

ή, πιο σύντομα,

$$\frac{\partial h_l}{\partial x_j}(\mathbf{x}_0) = \sum_{i=1}^m \frac{\partial g_l}{\partial y_i}(\mathbf{y}_0) \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0), \quad 1 \leq l \leq k, \quad 1 \leq j \leq n, \quad \mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0).$$

Ίσως βοηθήσει στο να θυμόμαστε καλύτερα την ισότητα (19) η κατανόηση της σχέσης ανάμεσα στις συντεταγμένες συναρτήσεις των $\mathbf{h}, \mathbf{g}, \mathbf{f}$. Από την $\mathbf{h} = \mathbf{g} \circ \mathbf{f}$ έχουμε

$$\mathbf{h}(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x})) \quad \text{για κάθε } \mathbf{x} \in A. \quad (20)$$

Έχουμε

$$\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{x}), \dots, h_k(\mathbf{x})) = (h_1(x_1, \dots, x_n), \dots, h_k(x_1, \dots, x_n)).$$

Επίσης,

$$\begin{aligned} \mathbf{g}(\mathbf{f}(\mathbf{x})) &= (g_1(\mathbf{f}(\mathbf{x})), \dots, g_k(\mathbf{f}(\mathbf{x}))) = (g_1(f_1(\mathbf{x}), \dots, f_m(\mathbf{x})), \dots, g_k(f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))) \\ &= (g_1(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)), \dots, g_k(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))). \end{aligned}$$

Επομένως, εξισώνοντας τις l -οστές συντεταγμένες των δυο πλευρών της (20) έχουμε

$$h_l(x_1, \dots, x_n) = g_l(f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)), \quad 1 \leq l \leq m.$$

Άρα μπορούμε να πούμε ότι η h_l προκύπτει ως σύνθεση της g_l με την m -άδα των $f_1, \dots, f_m$ σχηματικά ως εξής:

$$h_l(x_1, \dots, x_n) = g_l(y_1, \dots, y_m), \quad y_1 = f_1(x_1, \dots, x_n), \dots, y_m = f_m(x_1, \dots, x_n). \quad (21)$$

Τώρα, στην περίπτωση μίας μεταβλητής, από την σύνθεση

$$h(x) = g(y), \quad y = f(x)$$

προκύπτει η

$$\frac{dh}{dx}(x_0) = \frac{dg}{dy}(y_0) \frac{df}{dx}(x_0), \quad y_0 = f(x_0).$$

Το σημείο που πρέπει να προσέξουμε είναι ότι η g έχει μόνο μία μεταβλητή, την y , στην θέση της οποίας μπαίνει η μοναδική συντεταγμένη συνάρτηση f . Στην περίπτωση περισσότερων μεταβλητών, όπως αυτή εμφανίζεται στην (21), η g_l έχει $m \geq 2$ μεταβλητές, τις $y_1, \dots, y_m$, στις θέσεις των οποίων μπαίνουν οι m συντεταγμένες συναρτήσεις $f_1, \dots, f_m$. Έτσι, αντί να προκύψει ένα γινόμενο $\frac{dg}{dy}(y_0) \frac{df}{dx}(x_0)$, προκύπτουν m ανάλογα γινόμενα $\frac{\partial g_l}{\partial y_i}(\mathbf{y}_0) \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0)$, ένα για κάθε μεταβλητή y_i . Προσθέτουμε αυτά τα γινόμενα και έτσι βρίσκουμε το τελικό αποτέλεσμα (19) για την παράγωγο της h_l ως προς x_j .

Θα δούμε μερικά παραδείγματα. **Προσοχή:** Πρέπει να θυμόμαστε ότι όταν υπολογίζουμε παράγωγο συνάρτησης f μίας μεταβλητής x γράφουμε

$$\frac{df}{dx} \quad \text{αντί} \quad \frac{\partial f}{\partial x}.$$

1. Θεωρούμε τις $\mathbf{f} : \mathbb{R} \rightarrow \mathbb{R}^3$, $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ με τύπους

$$\mathbf{f}(x) = (x^2, 3x + 5, \sin x), \quad g(y, z, w) = yz^2 + w^2 e^z.$$

Τότε η σύνθεση $h = g \circ \mathbf{f} : \mathbb{R} \rightarrow \mathbb{R}$ έχει τύπο

$$h(x, y) = g(\mathbf{f}(x)) = g(x^2, 3x + 5, \sin x) = x^2(3x + 5)^2 + (\sin^2 x)e^{3x+5}.$$

Μπορούμε να υπολογίσουμε τις μερικές παραγώγους της h είτε από τον τελευταίο τύπο είτε μέσω του κανόνα της αλυσίδας, δηλαδή μέσω των τύπων (19). Ας εξασκηθούμε με τον δεύτερο τρόπο. Οι συντεταγμένες συναρτήσεις της $\mathbf{f}$ είναι οι

$$f_1(x) = x^2, \quad f_2(x) = 3x + 5, \quad f_3(x) = \sin x.$$

Όταν υπολογίζουμε τη σύνθεση αυτές οι συναρτήσεις μπαίνουν στη θέση των y, z, w αντιστοίχως. Άρα έχουμε

$$\begin{aligned} \frac{dh}{dx}(x) &= \frac{\partial g}{\partial y}(y, z, w) \frac{df_1}{dx}(x) + \frac{\partial g}{\partial z}(y, z, w) \frac{df_2}{dx}(x) + \frac{\partial g}{\partial w}(y, z, w) \frac{df_3}{dx}(x) \\ &= z^2 2x + (2yz + w^2 e^z) 3 + 2we^z \cos x \\ &= (3x + 5)^2 2x + (2x^2(3x + 5) + (\sin^2 x)e^{3x+5}) 3 + 2(\sin x)e^{3x+5} \cos x. \end{aligned}$$

2. Θεωρούμε τις $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ με τύπους

$$\mathbf{f}(x, y) = (x^2 y, x + e^y, \sin(xy)), \quad g(z, w, u) = z^2 w + w u^2.$$

Τότε η σύνθεση $h = g \circ \mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}$ έχει τύπο

$$h(x, y) = g(\mathbf{f}(x, y)) = g(x^2y, x + e^y, \sin(xy)) = x^4y^2(x + e^y) + (x + e^y) \sin^2(xy).$$

Οι συντεταγμένες συναρτήσεις της $\mathbf{f}$ είναι οι

$$f_1(x, y) = x^2y, \quad f_2(x, y) = x + e^y, \quad f_3(x, y) = \sin(xy).$$

Όταν υπολογίζουμε τη σύνθεση αυτές οι συναρτήσεις μπαίνουν στη θέση των z, w, u αντιστοίχως. Άρα έχουμε

$$\begin{aligned} \frac{\partial h}{\partial x}(x, y) &= \frac{\partial g}{\partial z}(z, w, u) \frac{\partial f_1}{\partial x}(x, y) + \frac{\partial g}{\partial w}(z, w, u) \frac{\partial f_2}{\partial x}(x, y) + \frac{\partial g}{\partial u}(z, w, u) \frac{\partial f_3}{\partial x}(x, y) \\ &= 2zw2xy + (z^2 + u^2) + 2wuy \cos(xy) \\ &= 2(x^2y)(x + e^y)2xy + (x^2y)^2 + \sin^2(xy) + 2(x + e^y) \sin(xy)y \cos(xy) \\ &= 4x^3y^2(x + e^y) + x^4y^2 + \sin^2(xy) + 2y(x + e^y) \sin(xy) \cos(xy) \end{aligned}$$

$$\begin{aligned} \frac{\partial h}{\partial y}(x, y) &= \frac{\partial g}{\partial z}(z, w, u) \frac{\partial f_1}{\partial y}(x, y) + \frac{\partial g}{\partial w}(z, w, u) \frac{\partial f_2}{\partial y}(x, y) + \frac{\partial g}{\partial u}(z, w, u) \frac{\partial f_3}{\partial y}(x, y) \\ &= 2zwx^2 + (z^2 + u^2)e^y + 2wux \cos(xy) \\ &= 2(x^2y)(x + e^y)x^2 + (x^2y)^2e^y + \sin^2(xy)e^y + 2(x + e^y) \sin(xy)x \cos(xy) \\ &= 2x^4y(x + e^y) + x^4y^2e^y + \sin^2(xy)e^y + 2x(x + e^y) \sin(xy) \cos(xy) \end{aligned}$$

3. Θεωρούμε τις $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^2$ με τύπους

$$f(x, y, z) = x^2ye^{x+z}, \quad \mathbf{g}(w) = (w^2, \sin w).$$

Τότε η σύνθεση $\mathbf{h} = \mathbf{g} \circ f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ έχει τύπο

$$\mathbf{h}(x, y, z) = \mathbf{g}(f(x, y, z)) = \mathbf{g}(x^2ye^{x+z}) = (x^4y^2e^{2x+2z}, \sin(x^2ye^{x+z})).$$

Η $\mathbf{g}$ έχει συντεταγμένες συναρτήσεις

$$g_1(w) = w^2, \quad g_2(w) = \sin w$$

και η $\mathbf{h}$ έχει συντεταγμένες συναρτήσεις

$$h_1(x, y, z) = g_1(f(x, y, z)), \quad h_2(x, y, z) = g_2(f(x, y, z)).$$

Υπολογίζοντας τις μερικές παραγώγους των h_1, h_2 μέσω των τύπων (19) βρίσκουμε

$$\begin{aligned} \frac{\partial h_1}{\partial x}(x, y, z) &= \frac{dg_1}{dw}(w) \frac{\partial f}{\partial x}(x, y, z) = 2w(2+x)xye^{x+z} = 2x^3y^2(2+x)e^{2x+2z} \\ \frac{\partial h_1}{\partial y}(x, y, z) &= \frac{dg_1}{dw}(w) \frac{\partial f}{\partial y}(x, y, z) = 2wx^2e^{x+z} = 2x^4ye^{2x+2z} \\ \frac{\partial h_1}{\partial z}(x, y, z) &= \frac{dg_1}{dw}(w) \frac{\partial f}{\partial z}(x, y, z) = 2wx^2ye^{x+z} = 2x^4y^2e^{2x+2z} \\ \frac{\partial h_2}{\partial x}(x, y, z) &= \frac{dg_2}{dw}(w) \frac{\partial f}{\partial x}(x, y, z) = (\cos w)(2+x)xye^{x+z} = \cos(x^2ye^{x+z})(2+x)xye^{x+z} \\ \frac{\partial h_2}{\partial y}(x, y, z) &= \frac{dg_2}{dw}(w) \frac{\partial f}{\partial y}(x, y, z) = (\cos w)x^2e^{x+z} = \cos(x^2ye^{x+z})x^2e^{x+z} \\ \frac{\partial h_2}{\partial z}(x, y, z) &= \frac{dg_2}{dw}(w) \frac{\partial f}{\partial z}(x, y, z) = (\cos w)x^2ye^{x+z} = \cos(x^2ye^{x+z})x^2ye^{x+z} \end{aligned}$$

8η εβδομάδα.

Μιλήσαμε για καμπύλες. **Καμπύλη** στον χώρο $\mathbb{R}^n$ είναι, αυστηρά μιλώντας, μία *συνάρτηση*

$$\sigma : I \rightarrow \mathbb{R}^n$$

όπου I είναι οποιοδήποτε διάστημα (όχι μονοσύνολο) του $\mathbb{R}$. Η μόνη προϋπόθεση για να λέμε ότι η συνάρτηση σ είναι καμπύλη είναι η συνέχειά της στο διάστημα I . Προσέξτε: το I πρέπει να είναι διάστημα και όχι ένωση διαστημάτων. Κατά τα άλλα, το I μπορεί να είναι διάστημα οποιουδήποτε τύπου: ανοικτό, κλειστό, ημιανοικτό, φραγμένο, μη-φραγμένο.

Αν $I = [a, b]$, τότε τα σημεία $\sigma(a)$, $\sigma(b)$ λέγονται **άκρα** της καμπύλης σ . Αν $I = (a, b)$, τότε λέμε ότι η καμπύλη σ δεν έχει άκρα. Αν το I περιέχει μόνο το ένα άκρο του, το a ή το b , τότε λέμε ότι η καμπύλη σ έχει μόνο ένα άκρο, το $\sigma(a)$ ή το $\sigma(b)$. Αν η καμπύλη έχει και τα δύο άκρα της και αυτά ταυτίζονται, δηλαδή αν $\sigma(a) = \sigma(b)$, τότε λέμε ότι η καμπύλη είναι **κλειστή**.

Συνήθως η ανεξάρτητη μεταβλητή της καμπύλης στο διάστημα I συμβολίζεται t . Έτσι θέλουμε να δηλώσουμε ότι όταν ο χρόνος t διατρέχει ένα χρονικό διάστημα I , το αντίστοιχο σημείο $\sigma(t)$ κινείται στον χώρο $\mathbb{R}^n$. Το σύνολο των θέσεων του κινητού σημείου $\sigma(t)$, δηλαδή το σύνολο τιμών της συνάρτησης σ είναι η λεγόμενη **τροχιά** της καμπύλης.

Καθώς η παράμετρος (χρόνος) t αυξάνεται στο διάστημα I το μεταβλητό σημείο $\sigma(t)$ κινείται πάνω στην τροχιά της καμπύλης σε μια συγκεκριμένη κατεύθυνση η οποία ονομάζεται **φορά** ή **κατεύθυνση** της καμπύλης ή και **φορά διαγραφής** της τροχιάς της καμπύλης.

Συνηθίζουμε να γράφουμε

$$\sigma(t) = (x_1(t), \dots, x_n(t)), \quad t \in I$$

όταν θέλουμε να δείξουμε τις συντεταγμένες συναρτήσεις της καμπύλης σ . Στην περίπτωση του $\mathbb{R}^2$ ή του $\mathbb{R}^3$ γράφουμε

$$\sigma(t) = (x(t), y(t)), \quad \sigma(t) = (x(t), y(t), z(t)).$$

Δύο απλά παραδείγματα.

1. Αν $\mathbf{x}_0$ είναι κάποιο σημείο στον $\mathbb{R}^n$ και $\mathbf{v}$ είναι κάποιο μη-μηδενικό διάνυσμα στον $\mathbb{R}^n$, τότε η συνάρτηση $\sigma : \mathbb{R} \rightarrow \mathbb{R}^n$ με τύπο

$$\sigma(t) = t\mathbf{v} + \mathbf{x}_0, \quad t \in \mathbb{R}$$

έχει ως τροχιά ολόκληρη την ευθεία η οποία είναι παράλληλη στο διάνυσμα $\mathbf{v}$ και διέρχεται από το σημείο $\mathbf{x}_0$. Η φορά διαγραφής της τροχιάς/ευθείας ταυτίζεται με την κατεύθυνση του $\mathbf{v}$.

Αν περιορίσουμε το πεδίο ορισμού της σ σε ένα υποδιάστημα I του $\mathbb{R}$, τότε η τροχιά της καμπύλης θα είναι ένα μέρος της ευθείας. Π.χ. αν $I = [a, b]$, τότε η τροχιά θα είναι το ευθύγραμμο τμήμα στον $\mathbb{R}^n$ με άκρα τα σημεία $\sigma(a) = a\mathbf{v} + \mathbf{x}_0$ και $\sigma(b) = b\mathbf{v} + \mathbf{x}_0$.

Είδαμε ότι αν μας δώσουν ένα οποιοδήποτε ευθύγραμμο τμήμα στον $\mathbb{R}^n$ με συγκεκριμένα άκρα $\mathbf{x}_0$ και $\mathbf{x}_1$, τότε μπορούμε να βρούμε μια καμπύλη η τροχιά της οποίας είναι το δοσμένο ευθύγραμμο τμήμα:

$$\sigma : [0, 1] \rightarrow \mathbb{R}^n, \quad \sigma(t) = t(\mathbf{x}_1 - \mathbf{x}_0) + \mathbf{x}_0 = t\mathbf{x}_1 + (1 - t)\mathbf{x}_0.$$

Η φορά διαγραφής του ευθυγράμμου τμήματος είναι από το άκρο $\sigma(0) = \mathbf{x}_0$ στο άκρο $\sigma(1) = \mathbf{x}_1$.

2. Αν (x_0, y_0) είναι σημείο του $\mathbb{R}^2$ και $r > 0$, τότε η συνάρτηση $\sigma : [0, 2\pi] \rightarrow \mathbb{R}^2$ με τύπο

$$\sigma(t) = (r \cos t + x_0, r \sin t + y_0), \quad t \in [0, 2\pi]$$

έχει ως τροχιά ολόκληρο τον κύκλο με κέντρο (x_0, y_0) και ακτίνα r : όταν το t διατρέχει το διάστημα $[0, 2\pi]$ το σημείο $\sigma(t)$ διατρέχει ολόκληρο το σύνολο $\{(x, y) \mid (x - x_0)^2 + (y - y_0)^2 = r^2\}$. Η συγκεκριμένη καμπύλη είναι προφανώς κλειστή: $\sigma(0) = \sigma(2\pi) = (r + x_0, y_0)$. Η φορά διαγραφής του κύκλου γύρω από το κέντρο του ταυτίζεται με την αντίθετη της φοράς περιστροφής των δεικτών του ρολογιού. Αυτή η φορά περιστροφής ονομάζεται **θετική φορά περιστροφής**.

Αν περιορίσουμε το πεδίο ορισμού της σ σε ένα υποδιάστημα I του $[0, 2\pi]$, τότε η τροχιά της καμπύλης θα είναι ένα τόξο του κύκλου.

Εδώ πρέπει να τονιστεί ένα σημείο που μπορεί να προκαλεί ασάφεια. Διακρίναμε ανάμεσα στην έννοια της καμπύλης (η οποία είναι μία συνάρτηση) και στην έννοια της τροχιάς της καμπύλης (το σύνολο τιμών της). Π.χ. στο δεύτερο παράδειγμα η καμπύλη είναι η συνάρτηση σ με τύπο $\sigma(t) = (r \cos t + x_0, r \sin t + y_0)$ ενώ η τροχιά της είναι ο κύκλος με κέντρο (x_0, y_0) και ακτίνα r , δηλαδή ένα σχήμα στο xy -επίπεδο. Όμως, έχουμε συνηθίσει παραδοσιακά να χαρακτηρίζουμε καμπύλη και τον κύκλο, δηλαδή το συγκεκριμένο σχήμα στο xy -επίπεδο. Το ίδιο γίνεται και στο πρώτο παράδειγμα. Χαρακτηρίζουμε -αυστηρά- καμπύλη την συνάρτηση σ με τύπο $\sigma(t) = t\mathbf{v} + \mathbf{x}_0$ και τροχιά της την ευθεία η οποία είναι παράλληλη στο διάνυσμα $\mathbf{v}$ και διέρχεται από το σημείο $\mathbf{x}_0$, αλλά παραδοσιακά χαρακτηρίζουμε καμπύλη και την ίδια την ευθεία. Δυστυχώς, αυτό είναι αναπόφευκτο και θα πρέπει κάθε φορά να ξεχωρίζουμε από τα συμφραζόμενα αν αναφερόμαστε σε καμπύλη-συνάρτηση ή σε καμπύλη-τροχιά.

Σύμφωνα με όσα έχουμε πει τις προηγούμενες εβδομάδες, η καμπύλη

$$\sigma : I \rightarrow \mathbb{R}^n$$

είναι παραγωγίσιμη ως διανυσματική συνάρτηση σε κάποιο εσωτερικό σημείο t_0 του I αν και μόνο αν καθεμία από τις συντεταγμένες συναρτήσεις $x_1, \dots, x_n$ είναι παραγωγίσιμη στο t_0 . Μάλιστα, σ' αυτήν την περίπτωση ο πίνακας ο οποίος ονομάζεται *παράγωγος* της σ στο t_0 είναι ο

$$D\sigma(t_0) = \begin{bmatrix} \frac{dx_1}{dt}(t_0) \\ \vdots \\ \frac{dx_n}{dt}(t_0) \end{bmatrix} = \begin{bmatrix} x'_1(t_0) \\ \vdots \\ x'_n(t_0) \end{bmatrix}$$

Κάναμε και τον εξής υπολογισμό:

$$\begin{aligned} \frac{\sigma(t) - \sigma(t_0)}{t - t_0} &= \frac{(x_1(t), \dots, x_n(t)) - (x_1(t_0), \dots, x_n(t_0))}{t - t_0} \\ &= \frac{(x_1(t) - x_1(t_0), \dots, x_n(t) - x_n(t_0))}{t - t_0} \\ &= \left(\frac{x_1(t) - x_1(t_0)}{t - t_0}, \dots, \frac{x_n(t) - x_n(t_0)}{t - t_0} \right). \end{aligned}$$

Άρα, όταν $t \rightarrow t_0$, το όριο του $\frac{\sigma(t) - \sigma(t_0)}{t - t_0}$ υπάρχει αν και μόνο αν υπάρχουν τα όρια όλων των $\frac{x_j(t) - x_j(t_0)}{t - t_0}$, δηλαδή όλες οι παράγωγοι $x'_j(t_0)$, και τότε έχουμε

$$\lim_{t \rightarrow t_0} \frac{\sigma(t) - \sigma(t_0)}{t - t_0} = \left(\lim_{t \rightarrow t_0} \frac{x_1(t) - x_1(t_0)}{t - t_0}, \dots, \lim_{t \rightarrow t_0} \frac{x_n(t) - x_n(t_0)}{t - t_0} \right) = (x'_1(t_0), \dots, x'_n(t_0)).$$

Το όριο $\lim_{t \rightarrow t_0} \frac{\sigma(t) - \sigma(t_0)}{t - t_0}$, αν υπάρχει, ονομάζεται κι αυτό **παράγωγος** της σ στο t_0 και συμβολίζεται $\sigma'(t_0)$. Δηλαδή:

$$\sigma'(t_0) = (x'_1(t_0), \dots, x'_n(t_0)).$$

Άρα η παράγωγος της καμπύλης είναι είτε ο $n \times 1$ πίνακας $D\sigma(t_0)$ είτε το n -διάστατο διάνυσμα $\sigma'(t_0)$. Και στις δύο περιπτώσεις οι συντεταγμένες της παραγώγου είναι οι παράγωγοι των συντεταγμένων συναρτήσεων.

Το διάνυσμα $\sigma'(t_0)$ ονομάζεται και (διανυσματική) **ταχύτητα** της καμπύλης, διότι εκφράζει τον στιγμιαίο ρυθμό μεταβολής της θέσης του κινητού $\sigma(t)$ ως προς τον χρόνο t την χρονική στιγμή t_0 . Το μήκος $\|\sigma'(t_0)\|$ της ταχύτητας ονομάζεται **αριθμητική ή βαθμωτή ταχύτητα**.

Είδαμε πολύ προσεκτικά ότι η παράγωγος (ή ταχύτητα) $\sigma'(t_0)$, αν την δούμε σαν διάνυσμα με αρχή το σημείο $\sigma(t_0)$ της τροχιάς, είναι εφαπτόμενη στην τροχιά στο σημείο $\sigma(t_0)$ και η κατεύθυνση του διανύσματος $\sigma'(t_0)$ συμπίπτει με την κατεύθυνση διαγραφής της τροχιάς κοντά στο σημείο $\sigma(t_0)$.

Κατόπιν θεωρήσαμε μία *πραγματική* συνάρτηση

$$f : A \rightarrow \mathbb{R}, \quad A \subseteq \mathbb{R}^n$$

και ένα σημείο $\mathbf{x}_0$ εσωτερικό του πεδίου ορισμού A της f . Επίσης θεωρήσαμε και μία καμπύλη

$$\sigma : I \rightarrow A$$

έτσι ώστε για κάποιο εσωτερικό σημείο t_0 του διαστήματος I να ισχύει $\sigma(t_0) = \mathbf{x}_0$. Κατόπιν θεωρήσαμε την σύνθεση

$$h = f \circ \sigma : I \rightarrow \mathbb{R}.$$

Τώρα, αν η σ είναι παραγωγίσιμη στο t_0 και η f είναι παραγωγίσιμη (ως συνάρτηση n μεταβλητών) στο $\sigma(t_0) = \mathbf{x}_0$, τότε η πραγματική συνάρτηση μίας μεταβλητής h είναι παραγωγίσιμη στο t_0 και από τον κανόνα της αλυσίδας έχουμε

$$Dh(t_0) = Df(\mathbf{x}_0)D\sigma(t_0). \quad (22)$$

Μάλιστα, αν γράψουμε αναλυτικά βάσει των συντεταγμένων,

$$h(t) = f(\sigma(t)) = f(x_1(t), \dots, x_n(t))$$

τότε ο κανόνας της αλυσίδας, δηλαδή η σχέση (22), λέει ότι

$$\frac{dh}{dt}(t_0) = \frac{\partial f}{\partial x_1}(\mathbf{x}_0) \frac{dx_1}{dt}(t_0) + \dots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0) \frac{dx_n}{dt}(t_0) = \frac{\partial f}{\partial x_1}(\mathbf{x}_0) x'_1(t_0) + \dots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0) x'_n(t_0).$$

Αμέσως παρατηρούμε ότι η τελευταία σχέση γράφεται και

$$h'(t_0) = \nabla f(\mathbf{x}_0) \cdot \sigma'(t_0). \quad (23)$$

Αυτή είναι μία πολύ χρήσιμη σχέση.

Αμέσως έχουμε μία πολύ χρήσιμη ειδική περίπτωση. Θεωρούμε την *ευθύγραμμη καμπύλη* του παραδείγματος 1:

$$\sigma : I \rightarrow \mathbb{R}^n$$

με τύπο

$$\sigma(t) = t\mathbf{v} + \mathbf{x}_0, \quad t \in I,$$

όπου θεωρούμε ότι το διάστημα I περιέχει το χρονικό σημείο $t_0 = 0$ ως εσωτερικό του σημείου, οπότε

$$\sigma(0) = \mathbf{x}_0$$

και, επίσης, θεωρούμε το διάστημα I γύρω από το $t_0 = 0$ να είναι αρκετά μικρό ώστε το ευθύγραμμο τμήμα το οποίο αποτελεί την τροχιά της καμπύλης σ να περιέχεται στο πεδίο ορισμού A της f (ώστε να έχει νόημα η σύνθεση $h = f \circ \sigma$). Τότε η σύνθεση $h = f \circ \sigma : I \rightarrow \mathbb{R}$ έχει φυσικά τύπο

$$h(t) = f(\sigma(t)) = f(\mathbf{x}_0 + t\mathbf{v}), \quad t \in I.$$

Υπολογίζουμε:

$$\sigma'(0) = \lim_{t \rightarrow 0} \frac{\sigma(t) - \sigma(0)}{t} = \lim_{t \rightarrow 0} \frac{(t\mathbf{v} + \mathbf{x}_0) - \mathbf{x}_0}{t} = \lim_{t \rightarrow 0} \mathbf{v} = \mathbf{v}.$$

Άρα σ' αυτήν την ειδική περίπτωση η σχέση (23) γράφεται

$$h'(0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v}. \quad (24)$$

Τώρα, αν σκεφτούμε ότι $h'(0) = \lim_{t \rightarrow 0} \frac{h(t) - h(0)}{t}$, βλέπουμε ότι η (24) γράφεται

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} = \nabla f(\mathbf{x}_0) \cdot \mathbf{v}. \quad (25)$$

Το αριστερό μέρος της (25) ονομάζεται **παράγωγος** της f στο σημείο $\mathbf{x}_0$ και **στην κατεύθυνση** $\mathbf{v}$ και συμβολίζεται $D_{\mathbf{v}}f(\mathbf{x}_0)$. Δηλαδή,

$$D_{\mathbf{v}}f(\mathbf{x}_0) := \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t}.$$

Άρα η σχέση (25) μας λέει ότι όταν η f είναι παραγωγίσιμη στο $\mathbf{x}_0$ ισχύει

$$D_{\mathbf{v}}f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v}. \quad (26)$$

Οι -ισοδύναμες- σχέσεις (25) και (26) μάς λένε το εξής ιδιαίτερα σημαντικό.

Αν θεωρήσουμε τις τιμές της f στα σημεία της ευθείας η οποία διέρχεται από το σημείο $\mathbf{x}_0$ και έχει την κατεύθυνση του διανύσματος $\mathbf{v}$, τότε ο στιγμιαίος ρυθμός μεταβολής αυτών των τιμών της f στο σημείο $\mathbf{x}_0$ είναι ίσος με το εσωτερικό γινόμενο της κλίσης της f στο $\mathbf{x}_0$ και του διανύσματος $\mathbf{v}$.

Όταν εφαρμόζουμε τα προηγούμενα, συνήθως θεωρούμε διανύσματα $\mathbf{v}$ μήκους 1:

$$\|\mathbf{v}\| = 1.$$

Στην περίπτωση που $\|\mathbf{v}\| = 1$ και που το διάνυσμα $\nabla f(\mathbf{x}_0)$ είναι μη-μηδενικό μπορούμε να θεωρήσουμε και την γωνία $\theta \in [0, \pi]$ την οποία σχηματίζουν τα διανύσματα $\nabla f(\mathbf{x}_0)$ και $\mathbf{v}$, και τότε οι (25), (26) γράφονται

$$D_{\mathbf{v}}f(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0)}{t} = \|\nabla f(\mathbf{x}_0)\| \cos \theta.$$

Οι τιμές του $\cos \theta$ κυμαίνονται στο διάστημα $[-1, 1]$ και είναι $\cos \theta = 1$ όταν $\theta = 0$ και $\cos \theta = -1$ όταν $\theta = \pi$. Επομένως έχουμε αμέσως τα εξής, επίσης σημαντικά, συμπεράσματα.

Ο μέγιστος (δηλαδή, μέγιστος θετικός) στιγμιαίος ρυθμός μεταβολής των τιμών της f στο σημείο $\mathbf{x}_0$ πιάνεται όταν το $\mathbf{v}$ έχει την ίδια κατεύθυνση με το $\nabla f(\mathbf{x}_0)$ και είναι ίσος με $\|\nabla f(\mathbf{x}_0)\|$.

Ο ελάχιστος (δηλαδή, μέγιστος αρνητικός) στιγμιαίος ρυθμός μεταβολής των τιμών της f στο σημείο $\mathbf{x}_0$ πιάνεται όταν το $\mathbf{v}$ έχει την αντίθετη κατεύθυνση με το $\nabla f(\mathbf{x}_0)$ και είναι ίσος με $-\|\nabla f(\mathbf{x}_0)\|$.

Μηδενικός στιγμιαίος ρυθμός μεταβολής των τιμών της f στο σημείο $\mathbf{x}_0$ πιάνεται όταν το $\mathbf{v}$ και το $\nabla f(\mathbf{x}_0)$ είναι κάθετα.

Τώρα θεωρούμε και πάλι μια οποιαδήποτε πραγματική συνάρτηση

$$f : A \rightarrow \mathbb{R}, \quad A \subseteq \mathbb{R}^2$$

(προσέξτε: $n = 2$) και ένα σημείο (x_0, y_0) εσωτερικό του πεδίου ορισμού A της f και υποθέτουμε ότι η f είναι παραγωγίσιμη στο (x_0, y_0) .

Θα μας απασχολήσει το ισοσταθμικό σύνολο της f το οποίο περιέχει το σημείο (x_0, y_0) :

$$M = \{(x, y) \in A \mid f(x, y) = f(x_0, y_0)\}.$$

Ειδικότερα θα μας απασχολήσει το πρόβλημα της εύρεσης διανύσματος κάθετου στο M στο σημείο (x_0, y_0) .

Υποθέτουμε ότι η κλίση της f στο (x_0, y_0) δεν είναι το μηδενικό διάνυσμα:

$$\nabla f(x_0, y_0) \neq (0, 0).$$

Με αυτήν την υπόθεση, όπως θα δούμε αργότερα στο μάθημα, μπορούμε να αποδείξουμε ότι το μέρος του M το οποίο βρίσκεται κοντά στο (x_0, y_0) αποτελεί τροχιά μιας καμπύλης $\sigma : I \rightarrow \mathbb{R}^2$ έτσι ώστε να είναι

$$\sigma(t_0) = (x_0, y_0)$$

για κάποιο εσωτερικό σημείο t_0 του I , αλλά και ώστε να είναι

$$\sigma'(t_0) \neq (0, 0).$$

Θεωρώντας αυτήν την καμπύλη σ , παίρνουμε την σύνθεση $h = f \circ \sigma : I \rightarrow \mathbb{R}$ και παρατηρούμε ότι

$$h(t) = f(\sigma(t)) = f(x_0, y_0) \quad \text{για κάθε } t \in I$$

επειδή $\sigma(t) \in M$ για κάθε $t \in I$. Άρα η h είναι σταθερή και άρα $h'(t_0) = 0$, οπότε η σχέση (23) δίνει

$$\nabla f(x_0, y_0) \cdot \sigma'(t_0) = 0. \quad (27)$$

Το διάνυσμα $\sigma'(t_0)$ είναι εφαπτόμενο στην τροχιά της σ στο σημείο $\sigma(t_0) = (x_0, y_0)$ και, επειδή αυτή η τροχιά ταυτίζεται με το μέρος του ισοσταθμικού συνόλου M το οποίο βρίσκεται κοντά στο σημείο (x_0, y_0) , το ίδιο διάνυσμα είναι εφαπτόμενο στο M στο σημείο (x_0, y_0) . Τώρα, από την (27) έχουμε ότι το $\nabla f(x_0, y_0)$ είναι κάθετο στο $\sigma'(t_0)$ και άρα είναι κάθετο στο M στο σημείο (x_0, y_0) .

Καταλήγουμε στο εξής συμπέρασμα:

Αν $\nabla f(x_0, y_0) \neq (0, 0)$, τότε το διάνυσμα αυτό είναι κάθετο στο ισοσταθμικό σύνολο της f , το οποίο περιέχει το σημείο (x_0, y_0) , στο σημείο (x_0, y_0) .

Τώρα θεωρούμε μια οποιαδήποτε πραγματική συνάρτηση

$$f : A \rightarrow \mathbb{R}, \quad A \subseteq \mathbb{R}^3$$

(προσέξτε: τώρα $n = 3$) και ένα σημείο (x_0, y_0, z_0) εσωτερικό του πεδίου ορισμού A της f και υποθέτουμε ότι η f είναι παραγωγίσιμη στο (x_0, y_0, z_0) .

Και πάλι θα μας απασχολήσει το ισοσταθμικό σύνολο της f το οποίο περιέχει το σημείο (x_0, y_0, z_0) ,

$$M = \{(x, y, z) \in A \mid f(x, y, z) = f(x_0, y_0, z_0)\},$$

και, ειδικότερα, το πρόβλημα της εύρεσης διανύσματος κάθετου στο M στο σημείο (x_0, y_0, z_0) .

Υποθέτουμε ότι η κλίση της f στο (x_0, y_0, z_0) δεν είναι το μηδενικό διάνυσμα:

$$\nabla f(x_0, y_0, z_0) \neq (0, 0, 0).$$

Θα δούμε αργότερα στο μάθημα ότι μπορούμε να αποδείξουμε τα εξής. Το μέρος του M το οποίο βρίσκεται κοντά στο (x_0, y_0, z_0) αποτελεί μία επιφάνεια η οποία έχει εφαπτόμενο επίπεδο στο σημείο (x_0, y_0, z_0) και, επίσης, για κάθε διάνυσμα $\mathbf{v}$ το οποίο έχει αρχή το (x_0, y_0, z_0) και βρίσκεται πάνω στο εφαπτόμενο επίπεδο υπάρχει καμπύλη $\sigma : I \rightarrow \mathbb{R}^3$ έτσι ώστε να είναι

$$\sigma(t_0) = (x_0, y_0, z_0), \quad \sigma'(t_0) = \mathbf{v}.$$

για κάποιο εσωτερικό σημείο t_0 του I και έτσι ώστε η τροχιά της σ να περιέχεται στο ισοσταθμικό σύνολο M .

Θεωρώντας αυτήν την καμπύλη σ παίρνουμε την σύνθεση $h = f \circ \sigma : I \rightarrow \mathbb{R}$ και παρατηρούμε ότι

$$h(t) = f(\sigma(t)) = f(x_0, y_0, z_0) \quad \text{για κάθε } t \in I$$

επειδή $\sigma(t) \in M$ για κάθε $t \in I$. Άρα η h είναι σταθερή και άρα $h'(t_0) = 0$, οπότε η σχέση (23) δίνει

$$\nabla f(x_0, y_0, z_0) \cdot \mathbf{v} = \nabla f(x_0, y_0, z_0) \cdot \sigma'(t_0) = 0.$$

Βλέπουμε, λοιπόν, ότι το $\nabla f(x_0, y_0, z_0)$ είναι κάθετο σε κάθε διάνυσμα $\mathbf{v}$ το οποίο έχει αρχή το (x_0, y_0, z_0) και βρίσκεται πάνω στο εφαπτόμενο επίπεδο του ισοσταθμικού συνόλου M στο σημείο (x_0, y_0, z_0) . Άρα το $\nabla f(x_0, y_0, z_0)$ είναι κάθετο στο M στο σημείο (x_0, y_0, z_0) .

Καταλήγουμε στο εξής συμπέρασμα:

Αν $\nabla f(x_0, y_0, z_0) \neq (0, 0, 0)$, τότε το διάνυσμα αυτό είναι κάθετο στο ισοσταθμικό σύνολο της f , το οποίο περιέχει το σημείο (x_0, y_0, z_0) , στο σημείο (x_0, y_0, z_0) .

9η εβδομάδα.

Αποδεικνύουμε το κριτήριο ολοκληρωσιμότητας στην ειδική περίπτωση πραγματικής συνάρτησης δύο μεταβλητών. Η απόδειξη στην γενική περίπτωση είναι παρόμοια. Θεωρούμε, λοιπόν, πραγματική συνάρτηση $f(x, y)$ η οποία έχει μερικές παραγώγους $\frac{\partial f}{\partial x}(x, y)$ και $\frac{\partial f}{\partial y}(x, y)$ σε κάθε σημείο (x, y) το οποίο βρίσκεται αρκούντως κοντά στο σημείο (x_0, y_0) και υποθέτουμε ότι αυτές οι μερικές παράγωγοι είναι συνεχείς στο (x_0, y_0) . Κατόπιν, γράφουμε

$$f(x, y) - f(x_0, y_0) = f(x, y) - f(x_0, y) + f(x_0, y) - f(x_0, y_0). \quad (28)$$

Θεωρώντας την συνάρτηση $f(x, y)$ ως συνάρτηση μόνο της μεταβλητής x , έχουμε

$$f(x, y) - f(x_0, y) = \frac{\partial f}{\partial x}(\xi, y)(x - x_0) \quad \text{για κάποιο } \xi \text{ ανάμεσα στα } x_0, x. \quad (29)$$

Θεωρώντας την συνάρτηση $f(x_0, y)$ ως συνάρτηση μόνο της μεταβλητής y , έχουμε

$$f(x_0, y) - f(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, \eta)(y - y_0) \quad \text{για κάποιο } \eta \text{ ανάμεσα στα } y_0, y. \quad (30)$$

Από τις σχέσεις (28), (29), (30) έχουμε ότι

$$f(x, y) - f(x_0, y_0) = \frac{\partial f}{\partial x}(\xi, y)(x - x_0) + \frac{\partial f}{\partial y}(x_0, \eta)(y - y_0)$$

και άρα

$$\begin{aligned} f(x, y) - f(x_0, y_0) &= \left[\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \right] \\ &+ \left[\frac{\partial f}{\partial x}(\xi, y) - \frac{\partial f}{\partial x}(x_0, y_0) \right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, \eta) - \frac{\partial f}{\partial y}(x_0, y_0) \right](y - y_0) \end{aligned}$$

Από την τελευταία σχέση παίρνουμε μετά από εφαρμογή της τριγωνικής ανισότητας:

$$\begin{aligned} 0 &\leq \frac{|f(x, y) - f(x_0, y_0) - [\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)]|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \\ &\leq \left| \frac{\partial f}{\partial x}(\xi, y) - \frac{\partial f}{\partial x}(x_0, y_0) \right| \frac{|x - x_0|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \\ &\quad + \left| \frac{\partial f}{\partial y}(x_0, \eta) - \frac{\partial f}{\partial y}(x_0, y_0) \right| \frac{|y - y_0|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} \\ &\leq \left| \frac{\partial f}{\partial x}(\xi, y) - \frac{\partial f}{\partial x}(x_0, y_0) \right| + \left| \frac{\partial f}{\partial y}(x_0, \eta) - \frac{\partial f}{\partial y}(x_0, y_0) \right| \end{aligned} \quad (31)$$

Τώρα παρατηρούμε ότι όταν $(x, y) \rightarrow (x_0, y_0)$ συνεπάγεται $x \rightarrow x_0$ και $y \rightarrow y_0$ και, επειδή το ξ είναι ανάμεσα στα x_0, x και το η είναι ανάμεσα στα y_0, y , συνεπάγεται $\xi \rightarrow x_0$ και $\eta \rightarrow y_0$.

Επομένως, όταν $(x, y) \rightarrow (x_0, y_0)$ συνεπάγεται $(\xi, y) \rightarrow (x_0, y_0)$ και $(x_0, \eta) \rightarrow (x_0, y_0)$. Τώρα, επειδή οι μερικές παράγωγοι της f είναι συνεχείς στο (x_0, y_0) συμπεραίνουμε ότι

$$\frac{\partial f}{\partial x}(\xi, y) \rightarrow \frac{\partial f}{\partial x}(x_0, y_0), \quad \frac{\partial f}{\partial y}(x_0, \eta) \rightarrow \frac{\partial f}{\partial y}(x_0, y_0).$$

Τέλος, από την (31) και το κριτήριο παρεμβολής, καταλήγουμε στο ότι

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{|f(x, y) - f(x_0, y_0) - [\frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)]|}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0$$

και άρα η f είναι, βάσει του ορισμού, παραγωγίσιμη στο (x_0, y_0) .

Κατόπιν, θεωρούμε παραγώγους συνάρτησης πολλών μεταβλητών τάξης μεγαλύτερης του ένα. Όπως και τις συναρτήσεις μίας μεταβλητής, ενδέχεται να μπορούμε να παραγωγίσουμε μια συνάρτηση πολλών μεταβλητών περισσότερες από μία φορές ως προς κάποιες μεταβλητές της. Να μερικά παραδείγματα:

1. Για την $f(x, y) = x^2y + xy^3$ έχουμε τις μερικές παραγώγους πρώτης τάξης

$$\frac{\partial f}{\partial x} = 2xy + y^3 \quad \frac{\partial f}{\partial y} = x^2 + 3xy^2$$

και άρα τις μερικές παραγώγους δεύτερης τάξης

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) &= 2y & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) &= 2x + 3y^2 \\ \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) &= 2x + 3y^2 & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) &= 6xy \end{aligned}$$

2. Για την $f(x, y) = \sin(xy^2)$ έχουμε τις μερικές παραγώγους πρώτης τάξης

$$\frac{\partial f}{\partial x} = y^2 \cos(xy^2) \quad \frac{\partial f}{\partial y} = 2xy \cos(xy^2)$$

και άρα τις μερικές παραγώγους δεύτερης τάξης

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) &= -y^4 \sin(xy^2) & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) &= 2y \cos(xy^2) - 2xy^3 \sin(xy^2) \\ \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) &= 2y \cos(xy^2) - 2xy^3 \sin(xy^2) & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) &= 2x \cos(xy^2) - 4x^2y^2 \sin(xy^2) \end{aligned}$$

3. Για την $f(x, y, z) = x^2yz^3$ έχουμε τις μερικές παραγώγους πρώτης τάξης

$$\frac{\partial f}{\partial x} = 2xyz^3 \quad \frac{\partial f}{\partial y} = x^2z^3 \quad \frac{\partial f}{\partial z} = 3x^2yz^2$$

και άρα τις μερικές παραγώγους δεύτερης τάξης

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) &= 2yz^3 & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) &= 2xz^3 & \frac{\partial}{\partial z} \left(\frac{\partial f}{\partial x} \right) &= 6xyz^2 \\ \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) &= 2xz^3 & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) &= 0 & \frac{\partial}{\partial z} \left(\frac{\partial f}{\partial y} \right) &= 3x^2z^2 \\ \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial z} \right) &= 6xyz^2 & \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial z} \right) &= 3x^2z^2 & \frac{\partial}{\partial z} \left(\frac{\partial f}{\partial z} \right) &= 6x^2yz \end{aligned}$$

Στα τρία προηγούμενα παραδείγματα θα μπορούσαμε να συνεχίσουμε παραγωγίζοντας κάθε μερική παράγωγο δεύτερης τάξης ως προς καθεμία από τις μεταβλητές παίρνοντας έτσι τις μερικές παραγώγους τρίτης τάξης και ούτω καθ' εξής.

Αν, λοιπόν, έχουμε συνάρτηση $f(\mathbf{x}) = f(x_1, \dots, x_n)$ τότε (με την προϋπόθεση ότι μπορούμε να παραγωγίζουμε) υπάρχουν n μερικές παράγωγοι πρώτης τάξης:

$$\frac{\partial f}{\partial x_j}(\mathbf{x}), \quad 1 \leq j \leq n,$$

n^2 μερικές παράγωγοι δεύτερης τάξης:

$$\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (\mathbf{x}), \quad 1 \leq i, j \leq n,$$

n^3 μερικές παράγωγοι τρίτης τάξης:

$$\frac{\partial}{\partial x_k} \left(\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) \right) (\mathbf{x}), \quad 1 \leq k, i, j \leq n,$$

και, γενικότερα, n^k μερικές παράγωγοι k -οστής τάξης.

Για απλούστευση των συμβόλων χρησιμοποιούμε τις συντμήσεις

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \quad \text{αντί} \quad \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (\mathbf{x})$$

$$\frac{\partial^3 f}{\partial x_k \partial x_i \partial x_j}(\mathbf{x}) \quad \text{αντί} \quad \frac{\partial}{\partial x_k} \left(\frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) \right) (\mathbf{x})$$

και ούτω καθ' εξής.

Στην πραγματικότητα το πλήθος των μερικών παραγώγων τάξης $k \geq 2$ είναι μικρότερο από το αναφερόμενο n^k . Αν παρατηρήσουμε προσεκτικά τα τρία παραδείγματα θα δούμε ότι η αλλαγή της σειράς παραγωγίσης ως προς δύο διαφορετικές μεταβλητές δεν αλλάζει το αποτέλεσμα. Στα δύο πρώτα παραδείγματα έχουμε

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

και στο τρίτο παράδειγμα

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x}, \quad \frac{\partial^2 f}{\partial y \partial z} = \frac{\partial^2 f}{\partial z \partial y}.$$

Αυτό δεν είναι τυχαίο και ισχύει κάτω από ασθενείς υποθέσεις, όπως το ότι αρκεί οι αντίστοιχες μερικές παράγωγοι να είναι συνεχείς. Ας το αποδείξουμε για συνάρτηση δύο μεταβλητών $f(x, y)$. Πιο συγκεκριμένα, υποθέτουμε ότι η πραγματική συνάρτηση $f(x, y)$ έχει μερικές παραγώγους δεύτερης τάξης $\frac{\partial^2 f}{\partial x \partial y}(x, y)$ και $\frac{\partial^2 f}{\partial y \partial x}(x, y)$ σε κάθε σημείο (x, y) το οποίο βρίσκεται αρκούντως κοντά στο σημείο (x_0, y_0) και ότι αυτές οι μερικές παράγωγοι είναι συνεχείς στο (x_0, y_0) . Με αυτές τις υποθέσεις θα αποδείξουμε ότι

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0). \quad (32)$$

Πράγματι, θεωρούμε την παράσταση

$$R(x, y) = f(x, y) - f(x, y_0) - f(x_0, y) + f(x_0, y_0) \quad \text{για } (x, y) \text{ κοντά στο } (x_0, y_0).$$

Ορίζουμε την συνάρτηση μίας μεταβλητής

$$g(x) = f(x, y) - f(x, y_0)$$

και παρατηρούμε ότι

$$R(x, y) = g(x) - g(x_0) = \frac{dg}{dx}(\xi)(x - x_0) = \left(\frac{\partial f}{\partial x}(\xi, y) - \frac{\partial f}{\partial x}(\xi, y_0) \right) (x - x_0) \quad (33)$$

για κάποιο ξ ανάμεσα στα x_0, x . Ομοίως, ορίζουμε την συνάρτηση μίας μεταβλητής

$$h(y) = \frac{\partial f}{\partial x}(\xi, y),$$

οπότε από την (33) έχουμε

$$R(x, y) = (h(y) - h(y_0))(x - x_0) = \frac{dh}{dy}(\eta)(y - y_0)(x - x_0) = \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta)(x - x_0)(y - y_0)$$

για κάποιο η ανάμεσα στα y_0, y . Από αυτό προκύπτει ότι

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{R(x, y)}{(x - x_0)(y - y_0)} = \lim_{(x,y) \rightarrow (x_0,y_0)} \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0), \quad (34)$$

επειδή η $\frac{\partial^2 f}{\partial y \partial x}$ είναι συνεχής στο (x_0, y_0) και $(\xi, \eta) \rightarrow (x_0, y_0)$ όταν $(x, y) \rightarrow (x_0, y_0)$.

Επαναλαμβάνουμε την ίδια διαδικασία αλλά με μια μικρή παραλλαγή. Τώρα ορίζουμε την συνάρτηση μίας μεταβλητής

$$g(y) = f(x, y) - f(x_0, y)$$

και παρατηρούμε ότι

$$R(x, y) = g(y) - g(y_0) = \frac{dg}{dy}(\bar{\eta})(y - y_0) = \left(\frac{\partial f}{\partial y}(x, \bar{\eta}) - \frac{\partial f}{\partial y}(x_0, \bar{\eta}) \right)(y - y_0) \quad (35)$$

για κάποιο $\bar{\eta}$ ανάμεσα στα y_0, y . (Αυτό το $\bar{\eta}$ δεν είναι κατ' ανάγκη ίδιο με το προηγούμενο η .) Ομοίως, ορίζουμε την συνάρτηση μίας μεταβλητής

$$h(x) = \frac{\partial f}{\partial y}(x, \bar{\eta}),$$

οπότε από την (35) έχουμε

$$R(x, y) = (h(x) - h(x_0))(y - y_0) = \frac{dh}{dx}(\bar{\xi})(x - x_0)(y - y_0) = \frac{\partial^2 f}{\partial x \partial y}(\bar{\xi}, \bar{\eta})(x - x_0)(y - y_0)$$

για κάποιο $\bar{\xi}$ ανάμεσα στα x_0, x . (Αυτό το $\bar{\xi}$ δεν είναι κατ' ανάγκη ίδιο με το προηγούμενο ξ .) Από αυτό προκύπτει ότι

$$\lim_{(x,y) \rightarrow (x_0,y_0)} \frac{R(x, y)}{(x - x_0)(y - y_0)} = \lim_{(x,y) \rightarrow (x_0,y_0)} \frac{\partial^2 f}{\partial x \partial y}(\bar{\xi}, \bar{\eta}) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0), \quad (36)$$

επειδή η $\frac{\partial^2 f}{\partial x \partial y}$ είναι συνεχής στο (x_0, y_0) και $(\bar{\xi}, \bar{\eta}) \rightarrow (x_0, y_0)$ όταν $(x, y) \rightarrow (x_0, y_0)$.

Από τις (34), (36) καταλήγουμε στην (32) που θέλουμε να αποδείξουμε.

Θεωρούμε την πράξη του πολλαπλασιασμού $m \times n$ πίνακα A με $n \times 1$ διάνυσμα-στήλη:

$$A\mathbf{x} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix}$$

Το αποτέλεσμα είναι $m \times 1$ διάνυσμα-στήλη και μας ενδιαφέρει να βρούμε μια σχέση ανάμεσα στα μήκη των διανυσμάτων $A\mathbf{x}$ και $\mathbf{x}$. Η σχέση αυτή είναι η:

$$\|A\mathbf{x}\| \leq M\|\mathbf{x}\|, \quad \text{όπου } M = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right)^{1/2}. \quad (37)$$

Η σχέση (37) προκύπτει ως εξής. Ονομάζουμε $\mathbf{a}_i$ το διάνυσμα-γραμμή $(a_{i1}, \dots, a_{in})$ και παρατηρούμε ότι η i -οστή συντεταγμένη του διανύσματος $A\mathbf{x}$ είναι ίση με το εσωτερικό γινόμενο $\mathbf{a}_i \cdot \mathbf{x}$, οπότε κάνουμε τους παρακάτω υπολογισμούς, χρησιμοποιώντας κάποια στιγμή την ανισότητα του Cauchy:

$$\|A\mathbf{x}\|^2 = \sum_{i=1}^m (\mathbf{a}_i \cdot \mathbf{x})^2 \leq \sum_{i=1}^m \|\mathbf{a}_i\|^2 \|\mathbf{x}\|^2 = \left(\sum_{i=1}^m \|\mathbf{a}_i\|^2 \right) \|\mathbf{x}\|^2 = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 \right) \|\mathbf{x}\|^2 = M^2 \|\mathbf{x}\|^2.$$

Διαγράφοντας τα τετράγωνα καταλήγουμε στην (37).

Τώρα θυμόμαστε το αποτέλεσμα από τον απειροστικό λογισμό μίας μεταβλητής το οποίο αναφέρεται ως **ανάπτυγμα Taylor n -οστής τάξης**: αν μία συνάρτηση $g(t)$ είναι $n - 1$ φορές παραγωγίσιμη σε κάποιο διάστημα $[a, b]$ και n φορές παραγωγίσιμη στο (a, b) , τότε

$$g(b) = g(a) + \frac{g'(a)}{1!}(b-a) + \frac{g''(a)}{2!}(b-a)^2 + \dots + \frac{g^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} + \frac{g^{(n)}(\xi)}{n!}(b-a)^n$$

για κάποιο $\xi \in (a, b)$. Στην ειδική περίπτωση $n = 1$ έχουμε το ανάπτυγμα Taylor πρώτης τάξης:

$$g(b) = g(a) + g'(\xi)(b-a). \quad (38)$$

Αυτό είναι απλώς το περιεχόμενο του θεωρήματος μέσης τιμής. Στην περίπτωση $n = 2$ έχουμε το ανάπτυγμα Taylor δεύτερης τάξης:

$$g(b) = g(a) + g'(a)(b-a) + \frac{g''(\xi)}{2}(b-a)^2. \quad (39)$$

Τώρα θα δούμε και θα αποδείξουμε τα αντίστοιχα αναπτύγματα Taylor πρώτης και δεύτερης τάξης για πραγματικές συναρτήσεις n μεταβλητών. Υπάρχει, φυσικά και το γενικότερο ανάπτυγμα Taylor τάξης μεγαλύτερης του 2, αλλά δεν θα αναφερθούμε σ' αυτό.

Για το ανάπτυγμα Taylor πρώτης τάξης υποθέτουμε ότι η πραγματική συνάρτηση $f(\mathbf{x}) = f(x_1, \dots, x_n)$ έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε κάθε σημείο $\mathbf{x}$ αρκούντως κοντά στο $\mathbf{x}_0$. Το συμπέρασμα είναι ότι ισχύει:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0j}) + R(\mathbf{x}) \quad (40)$$

όπου $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{x}_0 = (x_{01}, \dots, x_{0n})$ και όπου η συνάρτηση $R(\mathbf{x})$ έχει την ιδιότητα:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{R(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = 0. \quad (41)$$

Αν το καλοσκεφτούμε, αυτό το αποτέλεσμα είναι άμεση συνέπεια της παραγωγισιμότητας της f στο $\mathbf{x}_0$, αλλά θα το αποδείξουμε χρησιμοποιώντας την (38). Ορίζουμε την συνάρτηση

$$R(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}_0) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0j}),$$

οπότε αυτομάτως ισχύει η (40) και άρα απομένει να αποδείξουμε ότι ισχύει το όριο (41). Από την f ορίζουμε την συνάρτηση μίας μεταβλητής με τύπο

$$g(t) = f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)), \quad t \in [0, 1].$$

Η g είναι σύνθεση της f και της συνάρτησης $\mathbf{h}(t) = \mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)$. Επειδή η f έχει συνεχείς μερικές παραγώγους πρώτης τάξης, είναι παραγωγίσιμη. Η $\mathbf{h}$ είναι προφανώς παραγωγίσιμη και άρα μπορούμε να εφαρμόσουμε τον κανόνα αλυσίδας. Ο τύπος της $\mathbf{h}$ γράφεται

$$\mathbf{h}(t) = (h_1(t), \dots, h_n(t)) = (x_{01} + t(x_1 - x_{01}), \dots, x_{0n} + t(x_n - x_{0n}))$$

και έχουμε ότι

$$\begin{aligned}
 g'(t) &= \frac{dg}{dt}(t) = \frac{\partial f}{\partial x_1}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \frac{dh_1}{dt}(t) + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \frac{dh_n}{dt}(t) \\
 &= \frac{\partial f}{\partial x_1}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(x_1 - x_{01}) + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(x_n - x_{0n}) \\
 &= \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(x_j - x_{0j}).
 \end{aligned} \tag{42}$$

Έχουμε $g(0) = f(\mathbf{x}_0)$ και $g(1) = f(\mathbf{x})$, οπότε από την (38) στο διάστημα $[0, 1]$ παίρνουμε:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0 + \xi(\mathbf{x} - \mathbf{x}_0))(x_j - x_{0j})$$

για κάποιο $\xi \in (0, 1)$. Συγκρίνοντας αυτόν τον τύπο με τον τύπο που ορίζει την $R(\mathbf{x})$ βρίσκουμε ότι

$$R(\mathbf{x}) = \sum_{j=1}^n \left[\frac{\partial f}{\partial x_j}(\mathbf{x}_0 + \xi(\mathbf{x} - \mathbf{x}_0)) - \frac{\partial f}{\partial x_j}(\mathbf{x}_0) \right] (x_j - x_{0j}).$$

Απλουστεύουμε, θέτοντας

$$a_j(\mathbf{x}) = \frac{\partial f}{\partial x_j}(\mathbf{x}_0 + \xi(\mathbf{x} - \mathbf{x}_0)) - \frac{\partial f}{\partial x_j}(\mathbf{x}_0), \quad 1 \leq j \leq n, \tag{43}$$

οπότε

$$R(\mathbf{x}) = \sum_{j=1}^n a_j(\mathbf{x})(x_j - x_{0j}) = \mathbf{a}(\mathbf{x}) \cdot (\mathbf{x} - \mathbf{x}_0),$$

όπου

$$\mathbf{a}(\mathbf{x}) = (a_1(\mathbf{x}), \dots, a_n(\mathbf{x})).$$

Συνεπάγεται

$$|R(\mathbf{x})| \leq \|\mathbf{a}(\mathbf{x})\| \|\mathbf{x} - \mathbf{x}_0\|$$

και άρα

$$0 \leq \frac{|R(\mathbf{x})|}{\|\mathbf{x} - \mathbf{x}_0\|} \leq \|\mathbf{a}(\mathbf{x})\| = (a_1(\mathbf{x})^2 + \cdots + a_n(\mathbf{x})^2)^{1/2}. \tag{44}$$

Από τους τύπους (43) παίρνουμε ότι $a_j(\mathbf{x}) \rightarrow 0$ όταν $\mathbf{x} \rightarrow \mathbf{x}_0$, επειδή οι μερικές παράγωγοι $\frac{\partial f}{\partial x_j}$ είναι συνεχείς και επειδή το $\mathbf{x}_0 + \xi(\mathbf{x} - \mathbf{x}_0)$ τείνει στο $\mathbf{x}_0$ όταν $\mathbf{x} \rightarrow \mathbf{x}_0$ (το ξ είναι στο διάστημα $[0, 1]$ και άρα είναι φραγμένο). Τέλος, εφαρμόζοντας το κριτήριο παρεμβολής στην (44) καταλήγουμε στην (41).

Για το ανάπτυγμα Taylor δεύτερης τάξης υποθέτουμε ότι η πραγματική συνάρτηση $f(\mathbf{x}) = f(x_1, \dots, x_n)$ έχει συνεχείς μερικές παραγώγους δεύτερης τάξης σε κάθε σημείο $\mathbf{x}$ αρκούντως κοντά στο $\mathbf{x}_0$. Το συμπέρασμα είναι ότι ισχύει:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0j}) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)(x_i - x_{0i})(x_j - x_{0j}) + R(\mathbf{x}) \tag{45}$$

όπου $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{x}_0 = (x_{01}, \dots, x_{0n})$ και όπου η συνάρτηση $R(\mathbf{x})$ έχει την ιδιότητα:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{R(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^2} = 0. \tag{46}$$

Μιμούμενοι την προηγούμενη απόδειξη, ορίζουμε την συνάρτηση

$$R(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{x}_0) - \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0j}) - \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)(x_i - x_{0i})(x_j - x_{0j}),$$

οπότε αυτομάτως ισχύει η (45) και απομένει να αποδείξουμε ότι ισχύει το όριο (46). Από την f ορίζουμε την ίδια όπως πριν συνάρτηση μίας μεταβλητής με τύπο

$$g(t) = f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)), \quad t \in [0, 1].$$

Έχουμε ήδη υπολογίσει στον τύπο (42) την πρώτη παράγωγο της g . Από εκεί έχουμε:

$$g'(t) = \sum_{j=1}^n \frac{d}{dt} \left[\frac{\partial f}{\partial x_j}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \right] (x_j - x_{0j}). \quad (47)$$

Τώρα προσέξτε. Ο τύπος (42) λέει το εξής:

$$\frac{d}{dt} [f(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))] = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(x_i - x_{0i}).$$

(Αλλάξαμε το σύμβολο του δείκτη από j σε i .) Εφαρμόζοντας αυτό το αποτέλεσμα βάζοντας την συνάρτηση $\frac{\partial f}{\partial x_j}$ στη θέση της f βρίσκουμε ότι

$$\frac{d}{dt} \left[\frac{\partial f}{\partial x_j}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)) \right] = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(x_i - x_{0i}).$$

Άρα η σχέση (47) γίνεται

$$\begin{aligned} g'(t) &= \sum_{j=1}^n \left(\sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(x_i - x_{0i}) \right) (x_j - x_{0j}) \\ &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0))(x_i - x_{0i})(x_j - x_{0j}) \end{aligned} \quad (48)$$

Τώρα έχουμε $g(0) = f(\mathbf{x}_0)$, $g(1) = f(\mathbf{x})$ και, από την (42), $g'(0) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0j})$. Άρα από την (39) στο διάστημα $[0, 1]$ παίρνουμε:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0j}) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + \xi(\mathbf{x} - \mathbf{x}_0))(x_i - x_{0i})(x_j - x_{0j})$$

για κάποιο $\xi \in (0, 1)$. Συγκρίνοντας αυτόν τον τύπο με τον τύπο που ορίζει την $R(\mathbf{x})$ βρίσκουμε ότι

$$R(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^n \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + \xi(\mathbf{x} - \mathbf{x}_0)) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) \right] (x_i - x_{0i})(x_j - x_{0j}).$$

Απλουστεύουμε, θέτοντας

$$a_{ij}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0 + \xi(\mathbf{x} - \mathbf{x}_0)) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0), \quad 1 \leq i, j \leq n, \quad (49)$$

οπότε (δείτε το σχόλιο που ακολουθεί λίγο παρακάτω)

$$R(\mathbf{x}) = \frac{1}{2} \sum_{i,j=1}^n a_{ij}(\mathbf{x})(x_i - x_{0i})(x_j - x_{0j}) = \frac{1}{2} [A(\mathbf{x})(\mathbf{x} - \mathbf{x}_0)] \cdot (\mathbf{x} - \mathbf{x}_0), \quad (50)$$

όπου

$$A(\mathbf{x}) = \begin{bmatrix} a_{11}(\mathbf{x}) & \dots & a_{1n}(\mathbf{x}) \\ \vdots & & \vdots \\ a_{n1}(\mathbf{x}) & \dots & a_{nn}(\mathbf{x}) \end{bmatrix}$$

Τώρα χρησιμοποιούμε την ανισότητα (37) μαζί με την ανισότητα Cauchy και βρίσκουμε

$$|R(\mathbf{x})| \leq \frac{1}{2} \|A(\mathbf{x})(\mathbf{x} - \mathbf{x}_0)\| \|\mathbf{x} - \mathbf{x}_0\| \leq \frac{1}{2} M(\mathbf{x}) \|\mathbf{x} - \mathbf{x}_0\|^2,$$

όπου

$$M(\mathbf{x}) = \left(\sum_{i,j=1}^n a_{ij}(\mathbf{x})^2 \right)^{1/2}.$$

Άρα

$$0 \leq \frac{|R(\mathbf{x})|}{\|\mathbf{x} - \mathbf{x}_0\|^2} \leq \frac{1}{2} M(\mathbf{x}) = \frac{1}{2} \left(\sum_{i,j=1}^n a_{ij}(\mathbf{x})^2 \right)^{1/2}. \quad (51)$$

Από τους τύπους (49) παίρνουμε ότι $a_{ij}(\mathbf{x}) \rightarrow 0$ όταν $\mathbf{x} \rightarrow \mathbf{x}_0$, επειδή οι μερικές παράγωγοι $\frac{\partial^2 f}{\partial x_i \partial x_j}$ είναι συνεχείς και επειδή το $\mathbf{x}_0 + \xi(\mathbf{x} - \mathbf{x}_0)$ τείνει στο $\mathbf{x}_0$ όταν $\mathbf{x} \rightarrow \mathbf{x}_0$. Εφαρμόζουμε το κριτήριο παρεμβολής στην (51) και καταλήγουμε στην (46).

Σχόλιο. Στην σχέση (50) χρησιμοποιήσαμε κάτι γενικότερο που θα φανεί χρήσιμο και αργότερα. Ένα διπλό άθροισμα της μορφής

$$\sum_{i,j=1}^n a_{ij} x_i x_j$$

μπορούμε να το γράψουμε σε ισοδύναμη μορφή χρησιμοποιώντας κατάλληλο διάνυσμα και κατάλληλο πίνακα. Θεωρούμε (όπως κάναμε και στην απόδειξη παραπάνω) τον πίνακα

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

και το διάνυσμα $\mathbf{x} = (x_1, \dots, x_n)$, το οποίο μπορούμε να το γράψουμε, ανάλογα με τον σκοπό μας, και ως $n \times 1$ πίνακα-στήλη αλλά και ως $1 \times n$ πίνακα-γραμμή. Τότε έχουμε

$$\sum_{i,j=1}^n a_{ij} x_i x_j = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j \right) x_i = (A\mathbf{x}) \cdot \mathbf{x}, \quad (52)$$

αφού η i -οστή συντεταγμένη του διανύσματος $A\mathbf{x}$ είναι ακριβώς το $\sum_{j=1}^n a_{ij} x_j$. Αυτήν ακριβώς την ισοδύναμη μορφή χρησιμοποιήσαμε στην (50). Επίσης, μπορούμε να γράψουμε

$$\sum_{i,j=1}^n a_{ij} x_i x_j = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad (53)$$

όπου το διπλό άθροισμα ισούται με ένα γινόμενο τριών πινάκων. Αν γράψουμε το $\mathbf{x}$ ως $n \times 1$ πίνακα-στήλη τότε η (53) γράφεται

$$\sum_{i,j=1}^n a_{ij} x_i x_j = \mathbf{x}^\top A \mathbf{x}, \quad (54)$$

όπου, γενικά, με $B^\top$ συμβολίζουμε τον **συμμετρικό** πίνακα ενός πίνακα B . Αν γράψουμε το $\mathbf{x}$ ως $1 \times n$ πίνακα-γραμμή τότε η (53) γράφεται

$$\sum_{i,j=1}^n a_{ij} x_i x_j = \mathbf{x} A \mathbf{x}^\top. \quad (55)$$

10η εβδομάδα.

Έστω πραγματική συνάρτηση $f : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^n$, και $\mathbf{x}_0 \in A$. Λέμε ότι το $\mathbf{x}_0$ είναι σημείο **ολικού μεγίστου (ελαχίστου)** της f αν ισχύει

$$f(\mathbf{x}) \leq (\geq) f(\mathbf{x}_0) \quad \text{για κάθε } \mathbf{x} \in A.$$

Επίσης, λέμε ότι το $\mathbf{x}_0$ είναι σημείο **τοπικού μεγίστου (ελαχίστου)** της f αν υπάρχει κάποιο $\delta > 0$ ώστε να ισχύει

$$f(\mathbf{x}) \leq (\geq) f(\mathbf{x}_0) \quad \text{για κάθε } \mathbf{x} \in A \text{ με } \|\mathbf{x} - \mathbf{x}_0\| < \delta.$$

Ψάχνοντας μεθόδους για να βρίσκουμε τα τοπικά ακρότατα μιας συνάρτησης πολλών μεταβλητών, θυμόμαστε τί κάνουμε στην περίπτωση συναρτήσεων μίας μεταβλητής. Αν μια συνάρτηση ορίζεται σε ένα διάστημα I του $\mathbb{R}$ και αν το x_0 είναι **εσωτερικό σημείο** του I , τότε μία **αναγκαία** συνθήκη για να είναι το x_0 σημείο τοπικού ακροτάτου της f είναι η εξής: αν η f παραγωγίζεται στο x_0 τότε είναι $f'(x_0) = 0$. Έχουμε και τα σχετικά παραδείγματα. Η συνάρτηση $f(x) = x^2$ έχει τοπικό (και μάλιστα, ολικό) ελάχιστο στο 0, είναι παραγωγίσιμη στο 0 και $f'(0) = 0$. Η συνάρτηση $f(x) = |x|$ έχει τοπικό (και μάλιστα, ολικό) ελάχιστο στο 0 και δεν είναι παραγωγίσιμη στο 0. Από την άλλη μεριά έχουμε και παραδείγματα συναρτήσεων που δείχνουν ότι η παραπάνω συνθήκη δεν είναι ικανή. Δηλαδή, το να μην είναι η f παραγωγίσιμη στο x_0 ή το να ισχύει $f'(x_0) = 0$ δεν συνεπάγεται ότι το x_0 είναι σημείο τοπικού ακροτάτου της f . Η συνάρτηση με τύπο $f(x) = 2x$ για $x \geq 0$ και $f(x) = x$ για $x \leq 0$ δεν είναι παραγωγίσιμη στο 0 και δεν έχει τοπικό ακρότατο στο 0. Επίσης, η συνάρτηση $f(x) = x^3$ είναι παραγωγίσιμη στο 0 και ισχύει $f'(0) = 0$ και δεν έχει τοπικό ακρότατο στο 0. Πρέπει να τονιστεί ότι σε όλα τα προηγούμενα υποθέτουμε ότι το x_0 είναι εσωτερικό σημείο του διαστήματος I .

Μια ανάλογη αναγκαία συνθήκη για σημείο τοπικού ακροτάτου έχουμε και στην περίπτωση συνάρτησης πολλών μεταβλητών, το λεγόμενο **κριτήριο πρώτης παραγώγου**.

Έστω $f : A \rightarrow \mathbb{R}^n$, όπου $A \subseteq \mathbb{R}^n$, και $\mathbf{x}_0$ εσωτερικό σημείο του A το οποίο είναι σημείο τοπικού ακροτάτου της f . Αν η f είναι παραγωγίσιμη στο $\mathbf{x}_0$, τότε είναι $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

Φυσικά, η ισότητα $\nabla f(\mathbf{x}_0) = \mathbf{0}$ είναι ισοδύναμη με τις n ταυτόχρονες ισότητες

$$\frac{\partial f}{\partial x_1}(\mathbf{x}_0) = \dots = \frac{\partial f}{\partial x_n}(\mathbf{x}_0) = 0.$$

Για να αποδείξουμε αυτό το αποτέλεσμα υποθέτουμε ότι το $\mathbf{x}_0$ είναι εσωτερικό σημείο του πεδίου ορισμού A και σημείο τοπικού μεγίστου της f . Τότε υπάρχει κάποιο $\delta > 0$ ώστε να ισχύει

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \text{όταν } \|\mathbf{x} - \mathbf{x}_0\| < \delta. \quad (56)$$

Τώρα παίρνουμε ένα τυχαίο διάνυσμα $\mathbf{v} \neq \mathbf{0}$ και θεωρούμε την πραγματική συνάρτηση μίας μεταβλητής

$$h(t) = f(\mathbf{x}_0 + t\mathbf{v}) \quad \text{για } -\delta/\|\mathbf{v}\| < t < \delta/\|\mathbf{v}\|.$$

Παρατηρούμε ότι, αν $-\delta/\|\mathbf{v}\| < t < \delta/\|\mathbf{v}\|$, το σημείο $\mathbf{x}_0 + t\mathbf{v}$ έχει απόσταση από το $\mathbf{x}_0$ ίση με $\|\mathbf{x}_0 + t\mathbf{v} - \mathbf{x}_0\| = \|t\mathbf{v}\| = |t|\|\mathbf{v}\| < \delta$ και άρα, βάσει της (56), ισχύει

$$h(t) = f(\mathbf{x}_0 + t\mathbf{v}) \leq f(\mathbf{x}_0) = h(0) \quad \text{όταν } -\delta/\|\mathbf{v}\| < t < \delta/\|\mathbf{v}\|.$$

Άρα το 0 είναι εσωτερικό σημείο μεγίστου της h στο διάστημα $(-\delta/\|\mathbf{v}\|, \delta/\|\mathbf{v}\|)$. Τώρα, αν η f είναι παραγωγίσιμη στο $\mathbf{x}_0$, έχουμε αποδείξει στην όγδοη εβδομάδα ότι η h είναι παραγωγίσιμη στο 0 και ότι $h'(0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v}$: δείτε την σχέση (24). Άρα, αν η f είναι παραγωγίσιμη στο $\mathbf{x}_0$ συνεπάγεται ότι η h είναι παραγωγίσιμη στο 0 και, επειδή το 0 είναι σημείο μεγίστου της h , ισχύει $h'(0) = 0$ και επομένως $\nabla f(\mathbf{x}_0) \cdot \mathbf{v} = 0$. Αυτό, όμως, ισχύει για το οποιοδήποτε διάνυσμα $\mathbf{v} \neq \mathbf{0}$. Επειδή ισχύει, προφανώς, και για το διάνυσμα $\mathbf{0}$, έχουμε ότι ισχύει για κάθε $\mathbf{v}$. Άρα ισχύει και για το $\mathbf{v} = \nabla f(\mathbf{x}_0)$. Δηλαδή, ισχύει $\|\nabla f(\mathbf{x}_0)\|^2 = \nabla f(\mathbf{x}_0) \cdot \nabla f(\mathbf{x}_0) = 0$ και καταλήγουμε στο ότι $\nabla f(\mathbf{x}_0) = \mathbf{0}$.

Η απόδειξη είναι παρόμοια στην περίπτωση που το $\mathbf{x}_0$ είναι σημείο τοπικού ελαχίστου της f .

Παράδειγμα. Η συνάρτηση $f(x, y) = x^2 + y^2$ έχει πεδίο ορισμού το $\mathbb{R}^2$ οπότε κάθε υποψήφιο σημείο τοπικού ακροτάτου είναι εσωτερικό σημείο του πεδίου ορισμού. Επίσης, η f έχει συνεχείς μερικές παραγώγους $\frac{\partial f}{\partial x} = 2x$ και $\frac{\partial f}{\partial y} = 2y$ σε κάθε σημείο του $\mathbb{R}^2$ και άρα είναι παραγωγίσιμη σε κάθε υποψήφιο σημείο τοπικού ακροτάτου. Άρα, αν το (x, y) είναι σημείο τοπικού ακροτάτου της f , τότε πρέπει να είναι

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0$$

ή, ισοδύναμα,

$$\begin{cases} 2x = 0 \\ 2y = 0 \end{cases}$$

Άρα το μοναδικό σημείο τοπικού ακροτάτου της f είναι το $(0, 0)$. Πράγματι, είναι αμέσως φανερό ότι ισχύει

$$f(0, 0) = 0 \leq x^2 + y^2 = f(x, y) \quad \text{για κάθε } (x, y)$$

και άρα το $(0, 0)$ είναι σημείο (ολικού) ελαχίστου της f .

Παράδειγμα. Η συνάρτηση $f(x, y) = |x| + |y|$ έχει πεδίο ορισμού το $\mathbb{R}^2$. Είναι προφανές ότι το σημείο $(0, 0)$ είναι σημείο (ολικού) ελαχίστου της f . Αλλά η f δεν είναι παραγωγίσιμη στο $(0, 0)$. Πράγματι, η f δεν έχει παραγώγους πρώτης τάξης στο $(0, 0)$ αφού ούτε η συνάρτηση $f(x, 0) = |x|$ παραγωγίζεται στο 0 ούτε η συνάρτηση $f(0, y) = |y|$ παραγωγίζεται στο 0.

Παράδειγμα. Η συνάρτηση $f(x, y) = x^2 - y^2$ έχει συνεχείς μερικές παραγώγους $\frac{\partial f}{\partial x} = 2x$ και $\frac{\partial f}{\partial y} = -2y$ σε κάθε σημείο του $\mathbb{R}^2$ και άρα είναι παραγωγίσιμη σε κάθε υποψήφιο σημείο τοπικού ακροτάτου. Παρατηρούμε ότι στο σημείο $(0, 0)$ είναι $\frac{\partial f}{\partial x}(0, 0) = \frac{\partial f}{\partial y}(0, 0) = 0$, αλλά το $(0, 0)$ δεν είναι σημείο τοπικού ακροτάτου της f . Πράγματι, υπάρχουν σημεία (x, y) όσο κοντά θέλουμε στο $(0, 0)$ στα οποία η f έχει τιμές μεγαλύτερες από την τιμή στο $(0, 0)$ και υπάρχουν σημεία (x, y) όσο κοντά θέλουμε στο $(0, 0)$ στα οποία η f έχει τιμές μικρότερες από την τιμή στο $(0, 0)$, αφού ισχύει $f(x, 0) = x^2 > 0 = f(0, 0)$ για κάθε $x \neq 0$ και $f(0, y) = -y^2 < 0 = f(0, 0)$ για κάθε $y \neq 0$.

Παράδειγμα. Η συνάρτηση $f(x, y) = |x| - |y|$ έχει πεδίο ορισμού το $\mathbb{R}^2$. Όπως στο προηγούμενο παράδειγμα, το σημείο $(0, 0)$ δεν είναι σημείο τοπικού ακροτάτου της f . Επίσης, όπως στο προπροηγούμενο παράδειγμα, η f δεν είναι παραγωγίσιμη στο $(0, 0)$.

Παράδειγμα. Η συνάρτηση $f(x, y) = (x^2 + y^2)e^{-x^2 - y^2}$ έχει πεδίο ορισμού το $\mathbb{R}^2$ οπότε κάθε υποψήφιο σημείο τοπικού ακροτάτου είναι εσωτερικό σημείο του πεδίου ορισμού. Επίσης, η f έχει συνεχείς μερικές παραγώγους

$$\frac{\partial f}{\partial x} = 2xe^{-x^2 - y^2} - 2x(x^2 + y^2)e^{-x^2 - y^2}, \quad \frac{\partial f}{\partial y} = 2ye^{-x^2 - y^2} - 2y(x^2 + y^2)e^{-x^2 - y^2}$$

σε κάθε σημείο του $\mathbb{R}^2$ και άρα είναι παραγωγίσιμη σε κάθε υποψήφιο σημείο τοπικού ακροτάτου. Άρα, αν το (x, y) είναι σημείο τοπικού ακροτάτου της f , τότε πρέπει να είναι

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0$$

ή, ισοδύναμα,

$$\begin{cases} 2x(1 - x^2 - y^2)e^{-x^2-y^2} = 0 \\ 2y(1 - x^2 - y^2)e^{-x^2-y^2} = 0 \end{cases}$$

Λύνοντας αυτό το σύστημα βλέπουμε ότι τα υποψήφια σημεία τοπικού ακροτάτου της f είναι το $(0, 0)$ καθώς και όλα τα σημεία (x, y) με $x^2 + y^2 = 1$.

Είναι φανερό ότι το $(0, 0)$ είναι σημείο (ολικού) ελαχίστου της f αφού ισχύει

$$f(x, y) = (x^2 + y^2)e^{-x^2-y^2} \geq 0 = f(0, 0) \quad \text{για κάθε } (x, y).$$

Παίρνουμε τώρα τυχαίο σημείο (x_0, y_0) με $x_0^2 + y_0^2 = 1$ και ελέγχουμε αν είναι σημείο τοπικού ακροτάτου της f . Γι αυτό εξετάζουμε την ανισότητα

$$f(x, y) \leq f(x_0, y_0).$$

Αυτή ισοδυναμεί με

$$(x^2 + y^2)e^{-x^2-y^2} \leq e^{-1}$$

κι αυτή με

$$x^2 + y^2 \leq e^{x^2+y^2-1}.$$

Αυτή η ανισότητα ισχύει για κάθε (x, y) αφού η $e^{t-1} \geq t$ ισχύει για κάθε t .

Άρα το (x_0, y_0) είναι σημείο (ολικού) μεγίστου της f .

Μια άλλη μέθοδος που χρησιμοποιείται όταν ψάχνουμε τα ακρότατα συνάρτησης μίας μεταβλητής είναι το Θεώρημα Μέγιστης-Ελάχιστης Τιμής: αν η f είναι συνεχής στο κλειστό και φραγμένο διάστημα I , τότε η f έχει μέγιστη τιμή και ελάχιστη τιμή στο I . Αυτό το θεώρημα εξασφαλίζει την ύπαρξη σημείων ολικού μεγίστου και ολικού ελαχίστου στο I , οπότε βρίσκουμε τα υποψήφια σημεία ακροτάτου τα οποία είναι εσωτερικά σημεία του I (χρησιμοποιώντας, για παράδειγμα, το κριτήριο πρώτης παραγώγου) και συγκρίνουμε τις τιμές της f σ' αυτά τα εσωτερικά σημεία με τις τιμές της στα συνοριακά σημεία του I (δηλαδή στα άκρα του).

Ένα ανάλογο **Θεώρημα Μέγιστης-Ελάχιστης Τιμής** ισχύει και για συναρτήσεις πολλών μεταβλητών.

Έστω $f : A \rightarrow \mathbb{R}^n$, όπου το A είναι κλειστό και φραγμένο υποσύνολο του $\mathbb{R}^n$. Τότε η f έχει μέγιστη τιμή και ελάχιστη τιμή στο A . Δηλαδή, υπάρχουν $\mathbf{x}_1, \mathbf{x}_2 \in A$ για τα οποία ισχύει $f(\mathbf{x}_1) \leq f(\mathbf{x}) \leq f(\mathbf{x}_2)$ για κάθε $\mathbf{x} \in A$.

Το θεώρημα αυτό δεν θα το αποδείξουμε. Θυμόμαστε, μόνο, ότι το να είναι το A κλειστό σημαίνει ότι περιέχει όλα τα συνοριακά του σημεία.

Παράδειγμα. Θεωρούμε την συνάρτηση $f(x, y) = x^2 + 3y^2$ στον κλειστό μοναδιαίο δίσκο A (δηλαδή τον κλειστό δίσκο με κέντρο το $(0, 0)$ και ακτίνα 1):

$$f : A \rightarrow \mathbb{R}, \quad A = \{(x, y) \mid x^2 + y^2 \leq 1\}.$$

Το σύνολο A είναι κλειστό και φραγμένο. Άρα η f έχει μέγιστη τιμή και ελάχιστη τιμή στο A . Θα βρούμε τα πιθανά σημεία ακροτάτου στο εσωτερικό του A και κατόπιν θα βρούμε τα πιθανά σημεία ακροτάτου στο σύνορο του A . Το εσωτερικό του A , δηλαδή το σύνολο των εσωτερικών σημείων του A , είναι ο ανοικτός δίσκος $B = \{(x, y) \mid x^2 + y^2 < 1\}$ ενώ το σύνορο του A , δηλαδή το σύνολο των συνοριακών σημείων του A , είναι ο κύκλος $T = \{(x, y) \mid x^2 + y^2 = 1\}$.

Κάθε υποψήφιο σημείο τοπικού ακροτάτου στον ανοικτό δίσκο B είναι εσωτερικό σημείο του B και, επίσης, η f έχει συνεχείς μερικές παραγώγους $\frac{\partial f}{\partial x} = 2x$ και $\frac{\partial f}{\partial y} = 6y$ σε κάθε σημείο του B και άρα είναι παραγωγίσιμη σε κάθε υποψήφιο σημείο τοπικού ακροτάτου στο B . Άρα, αν το (x, y) είναι σημείο τοπικού ακροτάτου της f στο B , τότε πρέπει να είναι

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial f}{\partial y}(x, y) = 0$$

ή, ισοδύναμα,

$$\begin{cases} 2x = 0 \\ 6y = 0 \end{cases}.$$

Άρα το μοναδικό σημείο τοπικού ακροτάτου της f στο B είναι το $(0, 0)$ και η τιμή της f στο $(0, 0)$ είναι

$$f(0, 0) = 0.$$

Τώρα θα εξετάσουμε την f στο σύνορο του A , δηλαδή στον μοναδιαίο κύκλο T . Μία μέθοδος (η οποία είναι χρήσιμη και σε άλλες τέτοιες περιπτώσεις) είναι να περιγράψουμε τα σημεία του T χρησιμοποιώντας μόνο μία μεταβλητή αντί των δύο μεταβλητών x, y . Αυτό είναι εύκολο: θυμόμαστε ότι μια παραμετρικοποίηση των σημείων του κύκλου T είναι η

$$x = \cos t, \quad y = \sin t \quad \text{για } t \in [0, 2\pi].$$

Έτσι, οι τιμές $f(x, y) = x^2 + 3y^2$ για $x^2 + y^2 = 1$ γράφονται, ισοδύναμα,

$$f(\cos t, \sin t) = \cos^2 t + 3\sin^2 t = 1 + 2\sin^2 t \quad \text{για } t \in [0, 2\pi].$$

Άρα, αντί να βρούμε την μέγιστη τιμή και την ελάχιστη τιμή της f στον κύκλο T , μπορούμε, ισοδύναμα, να βρούμε την μέγιστη τιμή και την ελάχιστη τιμή της συνάρτησης *μίας μεταβλητής* $h(t) = 1 + 2\sin^2 t$ στο διάστημα $[0, 2\pi]$. Αυτό είναι εύκολο. Έχουμε

$$h'(t) = 4 \sin t \cos t,$$

οπότε η h' μηδενίζεται στα εσωτερικά σημεία $\pm\pi/2$ και π του $[0, 2\pi]$. Οι τιμές της h στα σημεία αυτά καθώς και στα άκρα του $[0, 2\pi]$ είναι:

$$h(0) = h(\pi) = h(2\pi) = 1, \quad h(\pm\pi/2) = 3.$$

Τα σημεία $t = 0, \pi, 2\pi$ του $[0, 2\pi]$ αντιστοιχούν στα σημεία $(1, 0), (-1, 0)$ του κύκλου T και τα σημεία $t = \pm\pi/2$ του $[0, 2\pi]$ αντιστοιχούν στα σημεία $(0, 1), (0, -1)$ του κύκλου T . Άρα η ελάχιστη τιμή και η μέγιστη τιμή της f στο σύνορο T του A είναι:

$$f(\pm 1, 0) = 1, \quad f(0, \pm 1) = 3.$$

Συγκρίνοντας τις τιμές της f στο $(0, 0)$ και στα $(\pm 1, 0)$ και $(0, \pm 1)$, έχουμε ότι η ελάχιστη τιμή και η μέγιστη τιμή της f στο A είναι:

$$f(0, 0) = 0, \quad f(0, \pm 1) = 3.$$

Μία ακόμη μέθοδος εύρεσης τοπικών ακροτάτων συνάρτησης *μίας μεταβλητής* βασίζεται στο λεγόμενο *κριτήριο δεύτερης παραγώγου*: αν στο εσωτερικό σημείο x_0 του πεδίου ορισμού της f ισχύει $f'(x_0) = 0$ και $f''(x_0) > (<) 0$ τότε το x_0 είναι σημείο τοπικού ελαχίστου (μεγίστου) της f . Μάλιστα, στην περίπτωση αυτή το x_0 είναι σημείο *γνήσιου* τοπικού ακροτάτου: για κάθε x αρκούντως κοντά στο x_0 και $x \neq x_0$ ισχύει $f(x) > (<) f(x_0)$.

Θα αναπτύξουμε ένα ανάλογο κριτήριο δεύτερης παραγώγου για συναρτήσεις πολλών μεταβλητών, αλλά πρώτα θα δούμε κάποια πράγματα από την Γραμμική Άλγεβρα.

Θεωρούμε συναρτήσεις n μεταβλητών της μορφής

$$Q(\mathbf{x}) = Q(x_1, \dots, x_n) = \sum_{i,j=1}^n a_{ij} x_i x_j = \sum_{1 \leq i,j \leq n} a_{ij} x_i x_j,$$

όπου οι n^2 αριθμοί a_{ij} θεωρούνται σταθεροί και δοσμένοι. Κάθε τέτοια συνάρτηση Q ονομάζεται **τετραγωνική μορφή** n μεταβλητών και είναι πραγματική και ορισμένη σε ολόκληρο

τον $\mathbb{R}^n$. Οι αριθμοί a_{ij} ονομάζονται **συντελεστές** της τετραγωνικής μορφής Q . Αν θεωρήσουμε τον $n \times n$ πίνακα

$$A = [a_{ij}]_{i,j=1}^n = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

με στοιχεία τους αριθμούς a_{ij} , είναι προφανές ότι η τετραγωνική μορφή Q και ο πίνακας A αλληλοκαθορίζονται: κάθε τετραγωνική μορφή αντιστοιχεί σε έναν πίνακα και κάθε πίνακας αντιστοιχεί σε μία τετραγωνική μορφή.

Εμείς θα θεωρήσουμε ότι οι αριθμοί a_{ij} έχουν μία επιπλέον ιδιότητα, την **συμμετρία**:

$$a_{ij} = a_{ji}.$$

Τότε η τετραγωνική μορφή Q ονομάζεται **συμμετρική** και, επίσης, όπως γνωρίζουμε από την Γραμμική Άλγεβρα, ο πίνακας A ονομάζεται **συμμετρικός**.

Πέρα από την απλή σχέση ανάμεσα στην τετραγωνική μορφή Q και στον πίνακα A η οποία έχει να κάνει με την ισότητα των συντελεστών της Q και των στοιχείων του A , υπάρχει και μία ουσιαστικότερη (και χρησιμότερη) σχέση η οποία εκφράζεται από τις ισότητες (52) και (54) της προηγούμενης εβδομάδας. Τις ξαναγράφουμε:

$$\sum_{i,j=1}^n a_{ij} x_i x_j = (A\mathbf{x}) \cdot \mathbf{x} = \mathbf{x}^\top A\mathbf{x}. \quad (57)$$

Λόγω της συμμετρίας $a_{ij} = a_{ji}$, κάθε όρος στο άθροισμα $\sum_{i,j=1}^n a_{ij} x_i x_j$ με διαφορετικούς δείκτες i, j υπάρχει δύο φορές: $a_{ij} x_i x_j = a_{ji} x_j x_i$. Γι αυτό η τετραγωνική μορφή γράφεται και

$$Q(\mathbf{x}) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j,$$

ομαδοποιώντας ξεχωριστά τους “διαγώνιους” όρους με $i = j$ και ξεχωριστά τους “μη-διαγώνιους” όρους με $i \neq j$.

Δύο παραδείγματα:

$$2x^2 + 3xy - y^2 = \left(\begin{bmatrix} 2 & 3/2 \\ 3/2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \right) \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 3/2 \\ 3/2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{aligned} x^2 + 2y^2 - 3z^2 + 2xy - 5xz - yz &= \left(\begin{bmatrix} 1 & 1 & -5/2 \\ 1 & 2 & -1/2 \\ -5/2 & -1/2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} \\ &= \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 & 1 & -5/2 \\ 1 & 2 & -1/2 \\ -5/2 & -1/2 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \end{aligned}$$

Λέμε ότι η συμμετρική τετραγωνική μορφή Q είναι **θετικά (αρνητικά) ορισμένη** και ότι ο αντίστοιχος συμμετρικός πίνακας A είναι **θετικά (αρνητικά) ορισμένος** αν ισχύει

$$Q(\mathbf{x}) > (<) 0 \quad \text{για κάθε } \mathbf{x} \neq \mathbf{0}.$$

ή, ισοδύναμα,

$$(A\mathbf{x}) \cdot \mathbf{x} > (<) 0 \quad \text{για κάθε } \mathbf{x} \neq \mathbf{0}.$$

Προφανώς, για $\mathbf{x} = \mathbf{0}$ έχουμε $Q(\mathbf{0}) = (A\mathbf{0}) \cdot \mathbf{0} = 0$.

Για παράδειγμα, η τετραγωνική μορφή $x^2 + 4y^2$ είναι θετικά ορισμένη, αφού έχει μη-αρνητικές τιμές και μηδενίζεται μόνο για $x = y = 0$. Ομοίως η τετραγωνική μορφή $x^2 - xy + y^2$ είναι θετικά

ορισμένη. Πράγματι, $x^2 - xy + y^2 = (x - y/2)^2 + 3y^2/4$ έχει μη-αρνητικές τιμές και μηδενίζεται μόνο για $x = y = 0$. Ομοίως η τετραγωνική μορφή $-x^2 - y^2 - 2z^2$ είναι αρνητικά ορισμένη, αφού έχει μη-θετικές τιμές και μηδενίζεται μόνο για $x = y = z = 0$.

Είναι γνωστό από την Γραμμική Άλγεβρα ότι κάθε συμμετρικός $n \times n$ πίνακας A έχει ακριβώς n (πραγματικές) ιδιοτιμές $\lambda_1, \dots, \lambda_n$ (κάποιες μπορεί να επαναλαμβάνονται) και ότι

$$\det(A) = \lambda_1 \cdots \lambda_n.$$

Επίσης, ισχύει ότι ο A είναι θετικά (αρνητικά) ορισμένος αν και μόνο αν όλες οι ιδιοτιμές του είναι θετικές (αρνητικές). Αυτό το τελευταίο θα το θεωρήσουμε δεδομένο αν και πολλές φορές δεν αποδεικνύεται στα πρώτα μαθήματα Γραμμικής Άλγεβρας.

Σχετικά με τα προηγούμενα παραδείγματα έχουμε τα εξής αντίστοιχα. Ο πίνακας της τετραγωνικής μορφής $x^2 + 4y^2$ είναι ο $\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ και έχει τις θετικές ιδιοτιμές 1 και 4. Ο πίνακας της

τετραγωνικής μορφής $x^2 - xy + y^2$ είναι ο $\begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix}$ και έχει τις θετικές ιδιοτιμές $1/2$ και

$3/2$. Ο πίνακας της τετραγωνικής μορφής $-x^2 - y^2 - 2z^2$ είναι ο $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ και έχει τις

αρνητικές ιδιοτιμές $-1, -1$ και -2 .

Το επόμενο αποτέλεσμα παρέχει ένα απλό κριτήριο για να είναι θετικά ή αρνητικά ορισμένη μια τετραγωνική μορφή δύο μεταβλητών ή ένας 2×2 πίνακας.

Ο συμμετρικός πίνακας $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ είναι θετικά (αρνητικά) ορισμένος ή, ισοδύναμα, η τετραγωνική μορφή $ax^2 + 2bxy + cy^2$ είναι θετικά (αρνητικά) ορισμένη αν και μόνο αν $a > (<) 0$ και $ac - b^2 > 0$. Παρατηρούμε ότι η παράσταση $ac - b^2$ είναι η ορίζουσα του πίνακα και ότι, αν $ac - b^2 > 0$, τότε το $a > (<) 0$ είναι ισοδύναμο με το $c > (<) 0$ (αφού $ac > b^2 \geq 0$).

Ας δούμε πώς αποδεικνύεται το παραπάνω κριτήριο. Έστω ότι ισχύει $a > 0$ και $ac - b^2 > 0$. Τότε

$$ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \frac{ac - b^2}{a}y^2 \quad (58)$$

και εύκολα βλέπουμε ότι η παράσταση έχει μη-αρνητικές τιμές και ότι μηδενίζεται μόνο όταν $x = y = 0$. Άρα η τετραγωνική μορφή είναι θετικά ορισμένη. Αντιστρόφως, έστω ότι η τετραγωνική μορφή είναι θετικά ορισμένη. Τότε με $x = 1$ και $y = 0$ βρίσκουμε ότι $a > 0$ και κατόπιν γράφουμε πάλι την προηγούμενη ταυτότητα (58) και βλέπουμε ότι για $y = 1$ και $x = -\frac{b}{a}$ προκύπτει $\frac{ac - b^2}{a} > 0$ και άρα $ac - b^2 > 0$. Φυσικά, η περίπτωση του αρνητικά ορισμένου πίνακα είναι παρόμοια.

Το κριτήριο αυτό μας επιτρέπει να βλέπουμε αμέσως αν ένας 2×2 πίνακας είναι θετικά (αρνητικά) ορισμένος χωρίς να χρειάζεται να βρούμε τις ιδιοτιμές του. Για παράδειγμα, για τον πίνακα $\begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix}$, που είδαμε προηγουμένως μέσω των ιδιοτιμών του ότι είναι θετικά ορισμένος, έχουμε $a = 1 > 0$ και $ac - b^2 = 1 \cdot 1 - (-1/2)^2 = 3/4 > 0$ και συμπεραίνουμε πολύ πιο γρήγορα ότι είναι θετικά ορισμένος.

Υπάρχει κι ένα συμπλήρωμα στο προηγούμενο κριτήριο για 2×2 πίνακες.

Αν $ac - b^2 < 0$, τότε (εξαιρώντας το ίδιο το σημείο $(0, 0)$) η τετραγωνική μορφή $ax^2 + 2bxy + cy^2$ έχει θετικές τιμές σε όλα τα σημεία μιας ευθείας που διέρχεται από το $(0, 0)$ και έχει αρνητικές τιμές σε όλα τα σημεία μιας άλλης ευθείας που διέρχεται από το $(0, 0)$.

Αυτό αποδεικνύεται εύκολα χρησιμοποιώντας την (58) και διακρίνοντας περιπτώσεις ως προς το αν τα a, c είναι θετικά, αρνητικά ή μηδέν. Για παράδειγμα, αν $a > 0$, τότε από την (58) φαίνεται αμέσως ότι ισχύει $ax^2 + 2bxy + cy^2 = ax^2 > 0$ για κάθε $(x, y) \neq (0, 0)$ στον x -άξονα και $ax^2 + 2bxy + cy^2 = (ac - b^2)y^2/a < 0$ για κάθε $(x, y) \neq (0, 0)$ στην ευθεία με καρτεσιανή εξίσωση $x = -(b/a)y$. Δείτε προσεκτικά τις διάφορες άλλες περιπτώσεις εσείς.

Αν λ_1, λ_2 είναι οι ιδιοτιμές του πίνακα $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, τότε έχουμε $\lambda_1 \lambda_2 = \det(A) = ac - b^2$, οπότε το τελευταίο κριτήριο αναδιατυπώνεται ως εξής:

Αν οι ιδιοτιμές του $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ είναι ετερόσημες, τότε (εξαιρώντας το ίδιο το σημείο $(0, 0)$) η τετραγωνική μορφή $ax^2 + 2bxy + cy^2$ έχει θετικές τιμές σε όλα τα σημεία μιας ευθείας που διέρχεται από το $(0, 0)$ και έχει αρνητικές τιμές σε όλα τα σημεία μιας άλλης ευθείας που διέρχεται από το $(0, 0)$.

Στην γενική περίπτωση συμμετρικής τετραγωνικής μορφής n μεταβλητών Q και του αντίστοιχου συμμετρικού $n \times n$ πίνακα A έχουμε την εξής γενίκευση:

Η συμμετρική τετραγωνική μορφή Q είναι θετικά (αρνητικά) ορισμένη αν και μόνο αν όλες οι ιδιοτιμές του αντίστοιχου συμμετρικού πίνακα A είναι θετικές (αρνητικές). Αν κάποιες από τις ιδιοτιμές του A είναι ετερόσημες, τότε (εξαιρώντας το ίδιο το σημείο $\mathbf{0}$) η τετραγωνική μορφή Q έχει θετικές τιμές σε όλα τα σημεία $\mathbf{x}$ μιας ευθείας που διέρχεται από το $\mathbf{0}$ και έχει αρνητικές τιμές σε όλα τα σημεία $\mathbf{x}$ μιας άλλης ευθείας που διέρχεται από το $\mathbf{0}$.

Το πρώτο συμπέρασμα θα το θεωρήσουμε, όπως είπαμε πιο πριν, δεδομένο χωρίς απόδειξη. Αλλά το δεύτερο μπορούμε να το αποδείξουμε εύκολα. Έστω λ μια θετική ιδιοτιμή του A . Αν $\mathbf{x} \neq \mathbf{0}$ είναι οποιοδήποτε ιδιοδιάνυσμα της λ (αυτά τα ιδιοδιανύσματα μαζί με το $\mathbf{0}$ σχηματίζουν τουλάχιστον μία ευθεία), τότε έχουμε $A\mathbf{x} = \lambda\mathbf{x}$ και άρα

$$Q(\mathbf{x}) = (A\mathbf{x}) \cdot \mathbf{x} = (\lambda\mathbf{x}) \cdot \mathbf{x} = \lambda\|\mathbf{x}\|^2 > 0.$$

Τώρα, έστω λ μια αρνητική ιδιοτιμή του A . Αν $\mathbf{x} \neq \mathbf{0}$ είναι οποιοδήποτε ιδιοδιάνυσμα της λ (πάλι αυτά τα ιδιοδιανύσματα μαζί με το $\mathbf{0}$ σχηματίζουν τουλάχιστον μία ευθεία), τότε έχουμε $A\mathbf{x} = \lambda\mathbf{x}$ και άρα

$$Q(\mathbf{x}) = (A\mathbf{x}) \cdot \mathbf{x} = (\lambda\mathbf{x}) \cdot \mathbf{x} = \lambda\|\mathbf{x}\|^2 < 0.$$

Μετά από όλα αυτά τα εισαγωγικά από την Γραμμική Άλγεβρα μπορούμε να διατυπώσουμε το κριτήριο δεύτερης παραγώγου για συναρτήσεις πολλών μεταβλητών. Θεωρούμε πραγματική συνάρτηση $f : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^n$, και σημείο $\mathbf{x}_0$ εσωτερικό του A . Υποθέτουμε ότι η f έχει συνεχείς μερικές παραγώγους δεύτερης τάξης σε κάθε σημείο $\mathbf{x}$ το οποίο βρίσκεται αρκούντως κοντά στο $\mathbf{x}_0$ και θεωρούμε την τετραγωνική μορφή η οποία εμφανίζεται στο ανάπτυγμα Taylor δεύτερης τάξης της f στο $\mathbf{x}_0$ (δείτε την σχέση (45) της προηγούμενης εβδομάδας):

$$\sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)(x_i - x_{0i})(x_j - x_{0j}).$$

Θέτουμε

$$\mathbf{h} = \mathbf{x} - \mathbf{x}_0$$

και άρα $h_i = x_i - x_{0i}$ για κάθε $i = 1, \dots, n$ και ορίζουμε τετραγωνική μορφή, την οποία συμβολίζουμε $H(f; \mathbf{x}_0)$, με τύπο

$$H(f; \mathbf{x}_0)(\mathbf{h}) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)h_i h_j.$$

Οι συντελεστές της τετραγωνικής μορφής είναι οι μερικές παράγωγοι δεύτερης τάξης της f στο $\mathbf{x}_0$ και, όπως έχουμε αποδείξει, είναι συμμετρικοί: $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x}_0)$. Αυτή η τετραγωνική μορφή ονομάζεται **Εссιανή τετραγωνική μορφή** της f στο $\mathbf{x}_0$ και ο αντίστοιχος συμμετρικός πίνακας, που τον συμβολίζουμε με το ίδιο σύμβολο $H(f; \mathbf{x}_0)$, είναι ο

$$H(f; \mathbf{x}_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}_0) \end{bmatrix}$$

και ονομάζεται **Εσσιανός πίνακας** της f στο $\mathbf{x}_0$.

Φτάνουμε (επιτέλους!) στο **κριτήριο δεύτερης παραγώγου**.

Έστω $f : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^n$, και σημείο $\mathbf{x}_0$ εσωτερικό του A . Υποθέτουμε ότι η f έχει συνεχείς μερικές παραγώγους δεύτερης τάξης σε κάθε σημείο $\mathbf{x}$ το οποίο βρίσκεται αρκούντως κοντά στο $\mathbf{x}_0$. Υποθέτουμε, επίσης, ότι $\nabla f(\mathbf{x}_0) = \mathbf{0}$. Τότε:

(i) Αν η Εσσιανή τετραγωνική μορφή της f στο $\mathbf{x}_0$ είναι θετικά (αρνητικά) ορισμένη ή, ισοδύναμα, αν όλες οι ιδιοτιμές του Εσσιανού πίνακα της f στο $\mathbf{x}_0$ είναι θετικές (αρνητικές), τότε το $\mathbf{x}_0$ είναι σημείο γνήσιου τοπικού ελαχίστου (μεγίστου) της f .

(ii) Αν ο Εσσιανός πίνακας της f στο $\mathbf{x}_0$ έχει κάποιες ετερόσημες ιδιοτιμές, τότε το $\mathbf{x}_0$ είναι **σαγματικό σημείο** της f . Δηλαδή, όσο κοντά θέλουμε στο $\mathbf{x}_0$ υπάρχουν $\mathbf{x}$ με $f(\mathbf{x}) > f(\mathbf{x}_0)$ και $\mathbf{x}$ με $f(\mathbf{x}) < f(\mathbf{x}_0)$.

Πριν αποδείξουμε αυτό το θεώρημα ας δούμε δύο παραδείγματα.

Παράδειγμα: Έστω η

$$f(x, y) = xy + \frac{1}{x} + \frac{8}{y}.$$

Η συνάρτηση ορίζεται στο ανοικτό σύνολο το οποίο αποτελείται από τον $\mathbb{R}^2$ χωρίς τον x -άξονα και τον y -άξονα, δηλαδή στην ένωση των τεσσάρων ανοικτών τεταρτημορίων του xy -επιπέδου. Η f έχει συνεχείς μερικές παραγώγους πρώτης τάξης

$$\frac{\partial f}{\partial x} = y - \frac{1}{x^2}, \quad \frac{\partial f}{\partial y} = x - \frac{8}{y^2}$$

σε κάθε σημείο του πεδίου ορισμού της και άρα είναι παραγωγίσιμη σε κάθε σημείο του πεδίου ορισμού της. Επίσης, κάθε σημείο του πεδίου ορισμού της f είναι εσωτερικό του σημείου οπότε μπορούμε να εφαρμόσουμε το κριτήριο πρώτης παραγώγου για να βρούμε τα πιθανά σημεία τοπικού ακροτάτου της f . Αν το (x, y) είναι σημείο τοπικού ακροτάτου, τότε

$$\begin{cases} y - \frac{1}{x^2} = 0 \\ x - \frac{8}{y^2} = 0 \end{cases}$$

Λύνουμε το σύστημα και βρίσκουμε το $(1/2, 4)$ ως το μοναδικό πιθανό σημείο τοπικού ακροτάτου της f . Τώρα, επειδή η f έχει και μερικές παραγώγους δεύτερης τάξης

$$\frac{\partial^2 f}{\partial x^2} = \frac{2}{x^3}, \quad \frac{\partial^2 f}{\partial y^2} = \frac{16}{y^3}, \quad \frac{\partial^2 f}{\partial x \partial y} = 1$$

συνεχείς σε κάθε σημείο του πεδίου ορισμού της (και ειδικά στα σημεία που είναι κοντά στο $(1/2, 4)$), εφαρμόζουμε το κριτήριο δεύτερης παραγώγου. Στο σημείο $(1/2, 4)$ έχουμε $\frac{\partial^2 f}{\partial x^2} = 16$, $\frac{\partial^2 f}{\partial y^2} = 1/4$ και $\frac{\partial^2 f}{\partial x \partial y} = 1$, οπότε ο Εσσιανός πίνακας της f στο $(1/2, 4)$ είναι ο

$$\begin{bmatrix} 16 & 1 \\ 1 & 1/4 \end{bmatrix}$$

Επειδή $16 > 0$ και $16 \cdot 1/4 - 1^2 = 3 > 0$, ο Εσσιανός πίνακας είναι θετικά ορισμένος και άρα το $(1/2, 4)$ είναι σημείο γνήσιου τοπικού ελαχίστου της f .

Αν και δεν χρειάζεται να τις υπολογίσουμε, οι ιδιοτιμές του Εσσιανού πίνακα είναι οι αριθμοί $(65 \pm \sqrt{4033})/8$ και είναι και οι δύο θετικοί, οπότε φαίνεται κι έτσι ότι ο πίνακας είναι θετικά ορισμένος.

Παράδειγμα: Έστω η

$$f(x, y) = x^2 + y^2 + \frac{1}{x^2 y^2}.$$

Όπως και στο προηγούμενο παράδειγμα, η συνάρτηση ορίζεται στο ανοικτό σύνολο το οποίο αποτελείται από τον $\mathbb{R}^2$ χωρίς τον x -άξονα και τον y -άξονα, δηλαδή στην ένωση των τεσσάρων ανοικτών τεταρτημορίων του xy -επιπέδου. Η f έχει συνεχείς μερικές παραγώγους πρώτης τάξης

$$\frac{\partial f}{\partial x} = 2x - \frac{2}{x^3 y^2}, \quad \frac{\partial f}{\partial y} = 2y - \frac{2}{x^2 y^3}$$

σε κάθε σημείο του πεδίου ορισμού της και άρα είναι παραγωγίσιμη σε κάθε σημείο του πεδίου ορισμού της. Επίσης, κάθε σημείο του πεδίου ορισμού της f είναι εσωτερικό του σημείου οπότε εφαρμόζουμε το κριτήριο πρώτης παραγώγου. Αν το (x, y) είναι σημείο τοπικού ακροτάτου, τότε

$$\begin{cases} 2x - \frac{2}{x^3 y^2} = 0 \\ 2y - \frac{2}{x^2 y^3} = 0 \end{cases}$$

Λύνουμε το σύστημα και βρίσκουμε τα τέσσερα σημεία $(1, 1)$, $(1, -1)$, $(-1, 1)$, $(-1, -1)$ ως τα μοναδικά πιθανά σημεία τοπικού ακροτάτου της f . Επειδή η f έχει και μερικές παραγώγους δεύτερης τάξης

$$\frac{\partial^2 f}{\partial x^2} = 2 + \frac{6}{x^4 y^2}, \quad \frac{\partial^2 f}{\partial y^2} = 2 + \frac{6}{x^2 y^4}, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{4}{x^3 y^3}$$

συνεχείς σε κάθε σημείο του πεδίου ορισμού της (και ειδικά στα σημεία που είναι κοντά στα τέσσερα σημεία που βρήκαμε), εφαρμόζουμε το κριτήριο δεύτερης παραγώγου. Και στα τέσσερα σημεία έχουμε $\frac{\partial^2 f}{\partial x^2} = 8$ και $\frac{\partial^2 f}{\partial y^2} = 8$. Στα σημεία $(1, 1)$, $(-1, -1)$ έχουμε $\frac{\partial^2 f}{\partial x \partial y} = 4$ ενώ στα σημεία $(1, -1)$, $(-1, 1)$ έχουμε $\frac{\partial^2 f}{\partial x \partial y} = -4$. Άρα στα τέσσερα σημεία ο Εσσιανός πίνακας της f είναι ο

$$\begin{bmatrix} 8 & \pm 4 \\ \pm 4 & 8 \end{bmatrix}$$

Επειδή $8 > 0$ και $8^2 - (\pm 4)(\pm 4) = 48 > 0$, ο Εσσιανός πίνακας είναι θετικά ορισμένος και άρα και τα τέσσερα σημεία είναι σημεία γνήσιου τοπικού ελαχίστου της f .

Και οι δύο Εσσιανοί πίνακες που προκύπτουν έχουν ως ιδιοτιμές τους αριθμούς 12 και 4, οι οποίοι είναι θετικοί και οι δύο. Δεν χρειάζεται να το κάνουμε αυτό, αλλά φαίνεται κι έτσι ότι οι Εσσιανοί πίνακες είναι και οι δύο θετικά ορισμένοι.

Bonus 1. Θα αποδείξουμε ότι ένας συμμετρικός $n \times n$ πίνακας A είναι θετικά ορισμένος αν και μόνο αν όλες οι ιδιοτιμές του είναι θετικές.

Έστω $\lambda_1, \dots, \lambda_n$ οι ιδιοτιμές του A . Μαθαίνουμε στην Γραμμική Άλγεβρα ότι υπάρχουν αντίστοιχα ιδιοδιανύσματα $\mathbf{x}_1, \dots, \mathbf{x}_n$ τα οποία αποτελούν ορθοκανονική βάση του $\mathbb{R}^n$. Δηλαδή τα διανύσματα αυτά έχουν μήκος ίσο με 1 και είναι ανά δύο κάθετα.

Κατ' αρχάς υποθέτουμε ότι ο A είναι θετικά ορισμένος, δηλαδή ότι $(A\mathbf{x}) \cdot \mathbf{x} > 0$ για κάθε $\mathbf{x} \neq \mathbf{0}$. Συνεπάγεται ότι για κάθε $\mathbf{x}_j$ ισχύει $(A\mathbf{x}_j) \cdot \mathbf{x}_j > 0$. Επειδή $A\mathbf{x}_j = \lambda_j \mathbf{x}_j$, συνεπάγεται $(\lambda_j \mathbf{x}_j) \cdot \mathbf{x}_j > 0$, οπότε $\lambda_j \|\mathbf{x}_j\|^2 > 0$ και άρα $\lambda_j > 0$. Επομένως, όλες οι ιδιοτιμές είναι θετικές.

Αντιστρόφως, έστω ότι $\lambda_1, \dots, \lambda_n > 0$. Θεωρούμε τυχόν $\mathbf{x} \neq \mathbf{0}$ και τότε υπάρχουν $\kappa_1, \dots, \kappa_n \in \mathbb{R}$ όχι όλοι ίσοι με μηδέν ώστε $\mathbf{x} = \kappa_1 \mathbf{x}_1 + \dots + \kappa_n \mathbf{x}_n$. Τότε

$$\begin{aligned} (A\mathbf{x}) \cdot \mathbf{x} &= \left(\sum_{i=1}^n \kappa_i A\mathbf{x}_i \right) \cdot \left(\sum_{j=1}^n \kappa_j \mathbf{x}_j \right) \\ &= \left(\sum_{i=1}^n \kappa_i \lambda_i \mathbf{x}_i \right) \cdot \left(\sum_{j=1}^n \kappa_j \mathbf{x}_j \right) \\ &= \sum_{i,j=1}^n \lambda_i \kappa_i \kappa_j (\mathbf{x}_i \cdot \mathbf{x}_j) = \sum_{i=1}^n \lambda_i \kappa_i^2 > 0. \end{aligned} \tag{59}$$

Άρα ο A είναι θετικά ορισμένος.

Μπορούμε να αποδείξουμε κάτι λίγο πιο συγκεκριμένο: αν ο A είναι θετικά ορισμένος, τότε υπάρχει κάποια θετική σταθερά $c > 0$ ώστε να ισχύει

$$(A\mathbf{x}) \cdot \mathbf{x} \geq c\|\mathbf{x}\|^2 \quad \text{για κάθε } \mathbf{x}. \quad (60)$$

Για να το αποδείξουμε ορίζουμε c να είναι η μικρότερη από τις ιδιοτιμές $\lambda_1, \dots, \lambda_n$, οπότε είναι $c > 0$ και ισχύει $\lambda_i \geq c$ για κάθε i . Τώρα, συνεχίζοντας την (59) βρίσκουμε

$$(A\mathbf{x}) \cdot \mathbf{x} = \sum_{i=1}^n \lambda_i \kappa_i^2 \geq \sum_{i=1}^n c \kappa_i^2 = c \sum_{i=1}^n \kappa_i^2. \quad (61)$$

Όμως, ισχύει επίσης

$$\|\mathbf{x}\|^2 = \mathbf{x} \cdot \mathbf{x} = \left(\sum_{i=1}^n \kappa_i \mathbf{x}_i \right) \cdot \left(\sum_{j=1}^n \kappa_j \mathbf{x}_j \right) = \sum_{i,j=1}^n \kappa_i \kappa_j (\mathbf{x}_i \cdot \mathbf{x}_j) = \sum_{i=1}^n \kappa_i^2. \quad (62)$$

Από τις (61) και (62) συνεπάγεται $(A\mathbf{x}) \cdot \mathbf{x} \geq c\|\mathbf{x}\|^2$.

Το τελευταίο συμπέρασμα διατυπώνεται, ισοδύναμα, και ως εξής. Αν η συμμετρική τετραγωνική μορφή Q είναι θετικά ορισμένη, τότε υπάρχει κάποια θετική σταθερά $c > 0$ ώστε να ισχύει

$$Q(\mathbf{x}) \geq c\|\mathbf{x}\|^2 \quad \text{για κάθε } \mathbf{x}. \quad (63)$$

Υπάρχει και το εξής ανάλογο αποτέλεσμα. Αν η συμμετρική τετραγωνική μορφή Q είναι αρνητικά ορισμένη και A είναι ο αντίστοιχος αρνητικά ορισμένος πίνακας, τότε υπάρχει κάποια θετική σταθερά $c > 0$ ώστε να ισχύει

$$Q(\mathbf{x}) = (A\mathbf{x}) \cdot \mathbf{x} \leq -c\|\mathbf{x}\|^2 \quad \text{για κάθε } \mathbf{x}. \quad (64)$$

Bonus 2. Θα αναφέρουμε (χωρίς απόδειξη) ένα ακόμη κριτήριο για να αποφασίζουμε αν ένας συμμετρικός $n \times n$ πίνακας A είναι θετικά ορισμένος ή αρνητικά ορισμένος.

Έστω

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$$

και έστω οι διαγώνιοι υπο-πίνακες του A με σταθερή άνω-αριστερή κορυφή το a_{11} , δηλαδή οι

$$A_1 = [a_{11}], \quad A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad A_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

και ούτω καθ' εξής μέχρι και τον πίνακα A_n ο οποίος ταυτίζεται με τον A .

Τότε ο A είναι θετικά ορισμένος αν και μόνο αν όλες οι ορίζουσες $\det(A_1), \det(A_2), \dots, \det(A_n)$ είναι θετικές. Επίσης, ο A είναι αρνητικά ορισμένος αν και μόνο αν οι ίδιες ορίζουσες είναι εναλλάξ αρνητικές και θετικές, αρχίζοντας με $\det(A_1) < 0$.

Φυσικά, $\det(A_1) = a_{11}$. Επίσης, παρατηρούμε ότι για 2×2 πίνακες A αυτό το κριτήριο ταυτίζεται με το κριτήριο που αποδείξαμε.

11η εβδομάδα.

Μια απλή ιδιότητα κάθε τετραγωνικής μορφής είναι η ταυτότητα:

$$Q(\lambda \mathbf{x}) = \lambda^2 Q(\mathbf{x}) \quad \text{για κάθε } \mathbf{x} \in \mathbb{R}^n, \lambda \in \mathbb{R}. \quad (65)$$

Πράγματι, επειδή $\lambda \mathbf{x} = (\lambda x_1, \dots, \lambda x_n)$, έχουμε

$$Q(\lambda \mathbf{x}) = \sum_{i,j=1}^n a_{ij} \lambda x_i \lambda x_j = \lambda^2 \sum_{i,j=1}^n a_{ij} x_i x_j = \lambda^2 Q(\mathbf{x}).$$

Θα αποδείξουμε ότι αν η συμμετρική τετραγωνική μορφή Q είναι θετικά (αρνητικά) ορισμένη ή, ισοδύναμα, αν ο αντίστοιχος συμμετρικός πίνακας A είναι θετικά (αρνητικά) ορισμένος, τότε υπάρχει κάποια *θετική* σταθερά $c > 0$ ώστε να ισχύει

$$Q(\mathbf{x}) \geq c \|\mathbf{x}\|^2 \quad (Q(\mathbf{x}) \leq -c \|\mathbf{x}\|^2) \quad \text{για κάθε } \mathbf{x}. \quad (66)$$

(Αυτό αποδείχθηκε μόνο στις σημειώσεις τις προηγούμενης εβδομάδας: οι σχέσεις (63) και (64). Δεν το αποδείξαμε στις διαλέξεις.)

Υποθέτουμε ότι η Q είναι θετικά ορισμένη. Τότε για κάθε $\mathbf{x}$ με $\|\mathbf{x}\| = 1$ ισχύει $Q(\mathbf{x}) > 0$. Επειδή η Q είναι συνεχής στην μοναδιαία σφαίρα του $\mathbb{R}^n$ (δηλαδή, στο σύνολο όλων των $\mathbf{x}$ με $\|\mathbf{x}\| = 1$) και επειδή η μοναδιαία σφαίρα είναι κλειστό και φραγμένο σύνολο, από το Θεώρημα Μέγιστης/Ελάχιστης Τιμής συνεπάγεται ότι η Q έχει ελάχιστη τιμή στην μοναδιαία σφαίρα. Άρα υπάρχει $\mathbf{x}_0$ με $\|\mathbf{x}_0\| = 1$ ώστε να ισχύει

$$Q(\mathbf{x}) \geq Q(\mathbf{x}_0) > 0$$

για κάθε $\mathbf{x}$ με $\|\mathbf{x}\| = 1$. Αν θέσουμε $c = Q(\mathbf{x}_0) > 0$, τότε ισχύει $Q(\mathbf{x}) \geq c > 0$ για κάθε $\mathbf{x}$ με $\|\mathbf{x}\| = 1$. Τώρα, θεωρούμε οποιοδήποτε $\mathbf{x} \neq \mathbf{0}$. Τότε το $\frac{1}{\|\mathbf{x}\|} \mathbf{x}$ έχει μήκος ίσο με 1 και άρα $Q(\frac{1}{\|\mathbf{x}\|} \mathbf{x}) \geq c$. Επομένως, από την (65) συνεπάγεται

$$\frac{1}{\|\mathbf{x}\|^2} Q(\mathbf{x}) = Q\left(\frac{1}{\|\mathbf{x}\|} \mathbf{x}\right) \geq c$$

από όπου προκύπτει η (66).

Τώρα θα αποδείξουμε το κριτήριο δεύτερης παραγώγου. Ξαναγράψουμε *περιληπτικά* το κριτήριο.

Θεωρούμε $f : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^n$, και σημείο $\mathbf{x}_0$ εσωτερικό του A . Υποθέτουμε ότι η f έχει συνεχείς μερικές παραγώγους δεύτερης τάξης σε κάθε σημείο $\mathbf{x}$ το οποίο βρίσκεται σε μια περιοχή του $\mathbf{x}_0$ και ότι $\nabla f(\mathbf{x}_0) = \mathbf{0}$. Τότε:

(i) Αν ο Εσσιανός πίνακας $H(f, \mathbf{x}_0)$ είναι θετικά (αρνητικά) ορισμένος, τότε το $\mathbf{x}_0$ είναι σημείο γνήσιου τοπικού ελαχίστου (μεγίστου) της f .

(ii) Αν ο Εσσιανός πίνακας $H(f, \mathbf{x}_0)$ έχει κάποιες ετερόσημες ιδιοτιμές, τότε το $\mathbf{x}_0$ είναι **σαγματικό σημείο** της f . Δηλαδή, όσο κοντά θέλουμε στο $\mathbf{x}_0$ υπάρχουν $\mathbf{x}$ με $f(\mathbf{x}) > f(\mathbf{x}_0)$ και $\mathbf{x}$ με $f(\mathbf{x}) < f(\mathbf{x}_0)$.

Η απόδειξη του (i) στην περίπτωση θετικά ορισμένου Εσσιανού πίνακα (η απόδειξη όταν ο Εσσιανός πίνακας είναι αρνητικά ορισμένος είναι παρόμοια) είναι η εξής. Αν ο πίνακας $H(f, \mathbf{x}_0)$ είναι θετικά ορισμένος, τότε και η Εσσιανή τετραγωνική μορφή είναι θετικά ορισμένη, οπότε υπάρχει κάποια σταθερά $c > 0$ ώστε

$$H(f, \mathbf{x}_0)(\mathbf{x}) \geq c\|\mathbf{x}\|^2 \quad \text{για κάθε } \mathbf{x}. \quad (67)$$

Τώρα θυμόμαστε το ανάπτυγμα Taylor δεύτερης τάξης της f στο $\mathbf{x}_0$, δηλαδή τις σχέσεις (45) και (46):

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0)(x_j - x_{0j}) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)(x_i - x_{0i})(x_j - x_{0j}) + R(\mathbf{x})$$

όπου $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{x}_0 = (x_{01}, \dots, x_{0n})$ και όπου η συνάρτηση $R(\mathbf{x})$ έχει την ιδιότητα:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{R(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|^2} = 0.$$

Γράφουμε $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$, οπότε $h_i = x_i - x_{0i}$ για κάθε i και άρα οι σχέσεις αυτές γράφονται

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{x}_0)h_j + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)h_i h_j + R(\mathbf{x}_0 + \mathbf{h}) \quad (68)$$

και

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{R(\mathbf{x}_0 + \mathbf{h})}{\|\mathbf{h}\|^2} = 0. \quad (69)$$

Παρατηρούμε ότι ο δεύτερος όρος στην δεξιά μεριά της (68) είναι ίσος με $\nabla f(\mathbf{x}_0) \cdot \mathbf{h} = \mathbf{0} \cdot \mathbf{h} = 0$. Επίσης, το διπλό άθροισμα στον τρίτο όρο στην δεξιά μεριά της (68) είναι ίσο με $H(f, \mathbf{x}_0)(\mathbf{h})$ και άρα η (68) γίνεται

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \frac{1}{2} H(f, \mathbf{x}_0)(\mathbf{h}) + R(\mathbf{x}_0 + \mathbf{h}). \quad (70)$$

Λόγω της (69), υπάρχει κάποιο $\delta > 0$ ώστε:

$$\frac{|R(\mathbf{x}_0 + \mathbf{h})|}{\|\mathbf{h}\|^2} < \frac{c}{2} \quad \text{για κάθε } \mathbf{h} \text{ με } 0 < \|\mathbf{h}\| < \delta, \quad (71)$$

όπου c είναι η σταθερά στην (67). Τώρα, συνδυάζοντας τις (67), (70) και (71), βρίσκουμε

$$f(\mathbf{x}_0 + \mathbf{h}) > f(\mathbf{x}_0) + \frac{c}{2} \|\mathbf{h}\|^2 - \frac{c}{2} \|\mathbf{h}\|^2 = f(\mathbf{x}_0) \quad \text{για κάθε } \mathbf{h} \text{ με } 0 < \|\mathbf{h}\| < \delta.$$

Ισοδύναμα,

$$f(\mathbf{x}) > f(\mathbf{x}_0) \quad \text{για κάθε } \mathbf{x} \text{ με } 0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta.$$

Άρα το $\mathbf{x}_0$ είναι σημείο γνήσιου τοπικού ελαχίστου.

Τώρα θα δούμε την απόδειξη του (ii), την οποία δεν κάναμε στις διαλέξεις.

Έστω $\lambda_1 > 0$ και $\lambda_2 < 0$ δύο ετερόσημες ιδιοτιμές του Εσσιανού πίνακα $H(f, \mathbf{x}_0)$. Τότε παίρνουμε $c = \min\{\lambda_1, -\lambda_2\} > 0$. Θεωρούμε οποιοδήποτε ιδιοδιάνυσμα $\mathbf{h}_1$ της λ_1 και οποιοδήποτε ιδιοδιάνυσμα $\mathbf{h}_2$ της λ_2 με $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 1$. Τότε

$$(H(f, \mathbf{x}_0)\mathbf{h}_1) \cdot \mathbf{h}_1 = (\lambda_1 \mathbf{h}_1) \cdot \mathbf{h}_1 = \lambda_1 \|\mathbf{h}_1\|^2 = \lambda_1 \geq c$$

και

$$(H(f, \mathbf{x}_0)\mathbf{h}_2) \cdot \mathbf{h}_2 = (\lambda_2 \mathbf{h}_2) \cdot \mathbf{h}_2 = \lambda_2 \|\mathbf{h}_2\|^2 = \lambda_2 \leq -c.$$

Τότε με τη Εσσιανή τετραγωνική μορφή έχουμε $H(f, \mathbf{x}_0)(\mathbf{h}_1) \geq c$ και $H(f, \mathbf{x}_0)(\mathbf{h}_2) \leq -c$. Πάλι, λόγω της (69), υπάρχει κάποιο $\delta > 0$ ώστε να ισχύει η (71), όπου τώρα η c είναι η νέα σταθερά που έχουμε ορίσει.

Τότε για κάθε $t \in \mathbb{R}$ με $0 < |t| < \delta$ είναι

$$H(f, \mathbf{x}_0)(t\mathbf{h}_1) = t^2 H(f, \mathbf{x}_0)(\mathbf{h}_1) \geq ct^2$$

και $0 < \|t\mathbf{h}_1\| = |t| < \delta$ και άρα, από την (71),

$$|R(\mathbf{x}_0 + t\mathbf{h}_1)| < \frac{c}{2} \|t\mathbf{h}_1\|^2 = \frac{c}{2} t^2.$$

Συνδυάζοντας τις δύο αυτές σχέσεις με την (70) βρίσκουμε

$$f(\mathbf{x}_0 + t\mathbf{h}_1) > f(\mathbf{x}_0) + \frac{c}{2} t^2 - \frac{c}{2} t^2 = f(\mathbf{x}_0) \quad \text{για κάθε } t \in \mathbb{R} \text{ με } 0 < |t| < \delta.$$

Άρα όταν $t \rightarrow 0$ το σημείο $\mathbf{x} = \mathbf{x}_0 + t\mathbf{h}_1$ πλησιάζει όσο θέλουμε κοντά στο $\mathbf{x}_0$ και για κάθε τέτοιο σημείο έχουμε $f(\mathbf{x}) > f(\mathbf{x}_0)$.

Ομοίως, για κάθε $t \in \mathbb{R}$ με $0 < |t| < \delta$ είναι

$$H(f, \mathbf{x}_0)(t\mathbf{h}_2) = t^2 H(f, \mathbf{x}_0)(\mathbf{h}_2) \leq -ct^2$$

και $0 < \|t\mathbf{h}_2\| = |t| < \delta$ και άρα, από την (71),

$$|R(\mathbf{x}_0 + t\mathbf{h}_2)| < \frac{c}{2} \|t\mathbf{h}_2\|^2 = \frac{c}{2} t^2.$$

Συνδυάζοντας αυτές τις δύο σχέσεις με την (70) βρίσκουμε

$$f(\mathbf{x}_0 + t\mathbf{h}_2) < f(\mathbf{x}_0) - \frac{c}{2} t^2 + \frac{c}{2} t^2 = f(\mathbf{x}_0) \quad \text{για κάθε } t \in \mathbb{R} \text{ με } 0 < |t| < \delta.$$

Άρα όταν $t \rightarrow 0$ το σημείο $\mathbf{x} = \mathbf{x}_0 + t\mathbf{h}_2$ πλησιάζει όσο θέλουμε κοντά στο $\mathbf{x}_0$ και για κάθε τέτοιο σημείο έχουμε $f(\mathbf{x}) < f(\mathbf{x}_0)$.

Το επόμενο θέμα μας είναι η μεγιστοποίηση ή ελαχιστοποίηση μιας πραγματικής συνάρτησης υπό συνθήκες.

Παράδειγμα: Αναζητάμε την μέγιστη τιμή και την ελάχιστη τιμή της συνάρτησης $f(x, y) = x^2 + xy + y^2$ όταν $x^2 + y^2 = 1$.

Η σχέση $x^2 + y^2 = 1$ είναι η λεγόμενη συνθήκη ή περιορισμός. Δηλαδή, μελετάμε τις τιμές της f όταν την περιορίσουμε στο σύνολο των σημείων (x, y) που ικανοποιούν την συγκεκριμένη συνθήκη, δηλαδή στον κύκλο με κέντρο $(0, 0)$ και ακτίνα 1, και από αυτές τις τιμές της f θέλουμε να βρούμε την μέγιστη και την ελάχιστη. Παρατηρούμε ότι το συγκεκριμένο σύνολο είναι ένα από τα ισοσταθμικά σύνολα της συνάρτησης $g(x, y) = x^2 + y^2$. Το γενικό θεώρημα βάσει του οποίου λύνεται το συγκεκριμένο πρόβλημα καθώς και πολλά άλλα παρόμοια είναι το εξής.

Έστω $f, g : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^n$, έστω εσωτερικό σημείο $\mathbf{x}_0$ του A και έστω ότι οι f, g έχουν συνεχείς μερικές παραγώγους πρώτης τάξης σε μια περιοχή του $\mathbf{x}_0$ και $\nabla g(\mathbf{x}_0) \neq \mathbf{0}$. Θεωρούμε το ισοσταθμικό σύνολο

$$S = \{\mathbf{x} \in A \mid g(\mathbf{x}) = c\}$$

και υποθέτουμε ότι $\mathbf{x}_0 \in S$, δηλαδή ότι $g(\mathbf{x}_0) = c$. Υποθέτουμε ότι το $\mathbf{x}_0$ είναι σημείο τοπικού μεγίστου (ελαχίστου) της f περιορισμένης στο σύνολο S , δηλαδή ότι υπάρχει κάποιο $\delta > 0$ ώστε να ισχύει $f(\mathbf{x}) \leq (\geq) f(\mathbf{x}_0)$ για κάθε $\mathbf{x} \in S$ με $\|\mathbf{x} - \mathbf{x}_0\| < \delta$. Τότε υπάρχει αριθμός λ ώστε

$$\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0).$$

Αυτό το αποτέλεσμα περιγράφει την λεγόμενη μέθοδο πολλαπλασιαστών Lagrange: ο αριθμός λ ονομάζεται πολλαπλασιαστής Lagrange.

Ας δούμε πώς θα εφαρμόσουμε το θεώρημα στο προηγούμενο παράδειγμα.

Παράδειγμα: Αναζητάμε την μέγιστη τιμή και την ελάχιστη τιμή της συνάρτησης $f(x, y) = x^2 + xy + y^2$ όταν $x^2 + y^2 = 1$.

Θεωρούμε την $g(x, y) = x^2 + y^2$ και το ισοσταθμικό σύνολο

$$S = \{(x, y) \mid g(x, y) = 1\} = \{(x, y) \mid x^2 + y^2 = 1\}.$$

Οι συναρτήσεις f, g έχουν συνεχείς μερικές παραγώγους σε ολόκληρο το $\mathbb{R}^2$:

$$\frac{\partial f}{\partial x} = 2x + y, \quad \frac{\partial f}{\partial y} = x + 2y, \quad \frac{\partial g}{\partial x} = 2x, \quad \frac{\partial g}{\partial y} = 2y.$$

Επίσης, για κάθε $(x, y) \in S$ ισχύει

$$\nabla g(x, y) = (2x, 2y) \neq (0, 0).$$

Άρα, αν υποθέσουμε ότι το $(x, y) \in S$ είναι σημείο τοπικού ακροτάτου της f περιορισμένης στο σύνολο S , τότε υπάρχει αριθμός λ ώστε

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

ή, ισοδύναμα,

$$(2x + y, x + 2y) = \lambda(2x, 2y).$$

Συνδυάζοντας με την σχέση $x^2 + y^2 = 1$ την οποία ικανοποιεί κάθε $(x, y) \in S$, βρίσκουμε ότι το σημείο τοπικού ακροτάτου (x, y) είναι λύση του συστήματος.

$$\begin{cases} 2x + y = 2\lambda x \\ x + 2y = 2\lambda y \\ x^2 + y^2 = 1 \end{cases}$$

Το σύστημα των δύο πρώτων εξισώσεων γράφεται

$$\begin{cases} 2(1 - \lambda)x + y = 0 \\ x + 2(1 - \lambda)y = 0 \end{cases}$$

Αν η ορίζουσα του συστήματος είναι $\neq 0$, τότε έχουμε $(x, y) = (0, 0)$ και αυτό το ζευγάρι δεν ικανοποιεί την τρίτη εξίσωση του αρχικού συστήματος. Άρα πρέπει η ορίζουσα να είναι 0:

$$4(1 - \lambda)^2 - 1 = 0$$

και άρα $\lambda = \frac{1}{2}$ ή $\lambda = \frac{3}{2}$. Με $\lambda = \frac{3}{2}$ βρίσκουμε από το δεύτερο σύστημα ότι $y = x$ και με $\lambda = \frac{1}{2}$ βρίσκουμε $y = -x$. Συνδυάζοντας με την τρίτη εξίσωση του αρχικού συστήματος βρίσκουμε τα τέσσερα πιθανά σημεία τοπικού ακροτάτου:

$$(\pm\sqrt{2}/2, \pm\sqrt{2}/2), \quad (\pm\sqrt{2}/2, \mp\sqrt{2}/2).$$

Στα δύο πρώτα σημεία η f έχει τιμή $\frac{3}{2}$ και στα δύο άλλα σημεία η f έχει τιμή $\frac{1}{2}$.

Από το Θεώρημα Μέγιστης/Ελάχιστης Τιμής η συνεχής συνάρτηση f έχει μέγιστη τιμή και ελάχιστη τιμή στο κλειστό και φραγμένο σύνολο S . Άρα οι δύο τιμές, $\frac{3}{2}$ και $\frac{1}{2}$, είναι η μέγιστη τιμή και η ελάχιστη τιμή της f υπό την συνθήκη $x^2 + y^2 = 1$.

Τώρα θα διατυπώσουμε το εξής πρόβλημα:

Θεωρούμε $f : A \rightarrow \mathbb{R}$, όπου το $A \subseteq \mathbb{R}^n$ είναι κλειστό και φραγμένο και η f έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε περιοχή κάθε σημείου του A . Υποθέτουμε ότι το σύνολο των συνοριακών σημείων του A , το σύνολο του A , είναι το ισοσταθμικό σύνολο μιας πραγματικής συνάρτησης

g , η οποία έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε περιοχή κάθε σημείου του συνόρου του A και ισχύει $\nabla g(\mathbf{x}) \neq \mathbf{0}$ για κάθε $\mathbf{x}$ στο σύνορο του A . Επειδή η f είναι συνεχής στο κλειστό και φραγμένο A , έχει μέγιστη τιμή και ελάχιστη τιμή στο A . Θέλουμε να υπολογίσουμε τις δύο αυτές ακρότατες τιμές της f καθώς και τα σημεία του A στα οποία πιάνονται οι ακρότατες τιμές της f .

Για να λύσουμε το πρόβλημα, πρώτα βρίσκουμε όλα τα πιθανά σημεία τοπικού ακροτάτου της f το εσωτερικό του A χρησιμοποιώντας το κριτήριο πρώτης παραγώγου και κατόπιν βρίσκουμε όλα τα πιθανά σημεία τοπικού ακροτάτου της f περιορισμένης στο σύνορο του A χρησιμοποιώντας την μέθοδο των πολλαπλασιαστών Lagrange. Αφού βρούμε όλα τα πιθανά σημεία τοπικού ακροτάτου της f είτε στο εσωτερικό του A είτε στο σύνορο του A , υπολογίζουμε τις τιμές της f στα σημεία αυτά: είναι προφανές ότι η μεγαλύτερη είναι η μέγιστη τιμή και η μικρότερη είναι η ελάχιστη τιμή της f συνολικά στο A .

Παράδειγμα: Αναζητάμε την μέγιστη τιμή και την ελάχιστη τιμή της συνάρτησης $f(x, y) = x^2 - 4xy + y^2$ στο σύνολο $A = \{(x, y) \mid x^2 + y^2 \leq 1\}$.

Το A είναι κλειστό και φραγμένο σύνολο, οπότε η f , η οποία είναι συνεχής στο A , έχει μέγιστη τιμή και ελάχιστη τιμή στο A . Η f έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε περιοχή κάθε σημείου του A (και μάλιστα σε ολόκληρο το $\mathbb{R}^2$). Το σύνορο του A είναι ο κύκλος $\{(x, y) \mid x^2 + y^2 = 1\}$ και είναι ισοσταθμικό σύνολο της συνάρτησης $g(x, y) = x^2 + y^2$, η οποία έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε περιοχή κάθε σημείου του συνόρου του A (και μάλιστα σε ολόκληρο το $\mathbb{R}^2$) και ισχύει $\nabla g(x, y) = (2x, 2y) \neq (0, 0)$ σε κάθε (x, y) στο σύνορο του A .

Βρίσκουμε τα πιθανά σημεία (x, y) τοπικού ακροτάτου της f στο εσωτερικό του A , δηλαδή με $x^2 + y^2 < 1$, εφαρμόζοντας το κριτήριο πρώτης παραγώγου:

$$\nabla f(x, y) = (0, 0)$$

ή, ισοδύναμα,

$$\begin{cases} 2x - 4y = 0 \\ -4x + 2y = 0 \end{cases}$$

Η μοναδική λύση του συστήματος είναι το $(x, y) = (0, 0)$.

Κατόπιν βρίσκουμε τα σημεία (x, y) τοπικού ακροτάτου της f περιορισμένης στο σύνορο του A , δηλαδή με $x^2 + y^2 = 1$, με την μέθοδο των πολλαπλασιαστών Lagrange. Κάτι παρόμοιο έγινε ακριβώς στο προηγούμενο παράδειγμα. Αν υποθέσουμε ότι το (x, y) με $x^2 + y^2 = 1$ είναι σημείο τοπικού ακροτάτου της f περιορισμένης στο σύνορο του A , τότε υπάρχει αριθμός λ ώστε

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

ή, ισοδύναμα,

$$(2x - 4y, -4x + 2y) = \lambda(2x, 2y).$$

Συνδυάζοντας με την σχέση $x^2 + y^2 = 1$, βρίσκουμε ότι το σημείο τοπικού ακροτάτου (x, y) είναι λύση του συστήματος.

$$\begin{cases} 2x - 4y = 2\lambda x \\ -4x + 2y = 2\lambda y \\ x^2 + y^2 = 1 \end{cases}$$

Το σύστημα των δύο πρώτων εξισώσεων γράφεται

$$\begin{cases} (1 - \lambda)x - 2y = 0 \\ -2x + (1 - \lambda)y = 0 \end{cases}$$

Αν η ορίζουσα του συστήματος είναι $\neq 0$, τότε έχουμε $(x, y) = (0, 0)$ και αυτό το ζευγάρι δεν ικανοποιεί την τρίτη εξίσωση του αρχικού συστήματος. Άρα πρέπει η ορίζουσα να είναι 0:

$$(1 - \lambda)^2 - 4 = 0$$

και άρα $\lambda = -1$ ή $\lambda = 3$. Με $\lambda = -1$ βρίσκουμε από το δεύτερο σύστημα ότι $y = x$ και με $\lambda = 3$ βρίσκουμε $y = -x$. Συνδυάζοντας με την τρίτη εξίσωση του αρχικού συστήματος βρίσκουμε τα τέσσερα πιθανά σημεία τοπικού ακροτάτου στο σύνορο του A :

$$(\pm\sqrt{2}/2, \pm\sqrt{2}/2), \quad (\pm\sqrt{2}/2, \mp\sqrt{2}/2).$$

Άρα συνολικά στο A έχουμε πέντε πιθανά σημεία τοπικού ακροτάτου:

$$(0, 0), \quad (\pm\sqrt{2}/2, \pm\sqrt{2}/2), \quad (\pm\sqrt{2}/2, \mp\sqrt{2}/2).$$

Στο $(0, 0)$ η f έχει τιμή 0, στα δύο επόμενα σημεία η f έχει τιμή -1 και στα δύο τελευταία σημεία η f έχει τιμή 3.

Άρα οι τιμές -1 και 3 είναι η ελάχιστη τιμή και η μέγιστη τιμή της f στο A .

Παράδειγμα: Θέλουμε να βρούμε την μέγιστη τιμή και την ελάχιστη τιμή της $f(x, y) = x - y + z$ στο σύνολο $A = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq 1\}$.

Το A , η κλειστή μπάλα στον $\mathbb{R}^3$ με κέντρο το $(0, 0, 0)$ και ακτίνα 1, είναι κλειστό και φραγμένο σύνολο, οπότε η f , η οποία είναι συνεχής στο A , έχει μέγιστη τιμή και ελάχιστη τιμή στο A . Η f έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε περιοχή κάθε σημείου του A (και μάλιστα σε ολόκληρο το $\mathbb{R}^3$). Το σύνορο του A είναι η σφαίρα $\{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$ και είναι ισοσταθμικό σύνολο της συνάρτησης $g(x, y, z) = x^2 + y^2 + z^2$, η οποία έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε περιοχή κάθε σημείου του συνόρου του A (και μάλιστα σε ολόκληρο το $\mathbb{R}^3$) και ισχύει $\nabla g(x, y, z) = (2x, 2y, 2z) \neq (0, 0, 0)$ σε κάθε (x, y, z) στο σύνορο του A .

Βρίσκουμε τα πιθανά σημεία (x, y, z) τοπικού ακροτάτου της f στο εσωτερικό του A , δηλαδή με $x^2 + y^2 + z^2 < 1$, εφαρμόζοντας το κριτήριο πρώτης παραγώγου:

$$\nabla f(x, y, z) = (0, 0, 0)$$

ή, ισοδύναμα,

$$(1, -1, 1) = (0, 0, 0),$$

το οποίο είναι αδύνατο. Άρα δεν υπάρχει κανένα εσωτερικό σημείο τοπικού ακροτάτου.

Τώρα βρίσκουμε τα σημεία (x, y, z) τοπικού ακροτάτου της f περιορισμένης στο σύνορο του A , δηλαδή με $x^2 + y^2 + z^2 = 1$, με την μέθοδο των πολλαπλασιαστών Lagrange. Αν το (x, y, z) με $x^2 + y^2 + z^2 = 1$ είναι σημείο τοπικού ακροτάτου της f περιορισμένης στο σύνορο του A , τότε υπάρχει αριθμός λ ώστε

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$$

ή, ισοδύναμα,

$$(1, -1, 1) = \lambda(2x, 2y, 2z).$$

Συνδυάζοντας με την σχέση $x^2 + y^2 + z^2 = 1$, βρίσκουμε ότι το σημείο τοπικού ακροτάτου (x, y, z) είναι λύση του συστήματος.

$$\begin{cases} 1 = 2\lambda x \\ -1 = 2\lambda y \\ 1 = 2\lambda z \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

Προφανώς είναι $\lambda \neq 0$, οπότε από τις τρεις πρώτες εξισώσεις του συστήματος παίρνουμε

$$x = \frac{1}{2\lambda}, \quad y = -\frac{1}{2\lambda}, \quad z = \frac{1}{2\lambda}.$$

Τότε η τέταρτη εξίσωση γίνεται

$$\frac{3}{4\lambda^2} = 1$$

και άρα $\lambda = -\sqrt{3}/2$ ή $\lambda = \sqrt{3}/2$. Άρα έχουμε δύο πιθανά σημεία τοπικού ακροτάτου στο σύνορο του A :

$$(-1/\sqrt{3}, 1/\sqrt{3}, -1/\sqrt{3}), \quad (1/\sqrt{3}, -1/\sqrt{3}, 1/\sqrt{3}).$$

Στο πρώτο σημείο η f έχει τιμή $-\sqrt{3}$ και στο δεύτερο σημείο έχει τιμή $\sqrt{3}$. Άρα οι τιμές $-\sqrt{3}$ και $\sqrt{3}$ είναι η ελάχιστη τιμή και η μέγιστη τιμή της f στο A .

Τώρα θα περιγράψουμε την μέθοδο των πολλαπλασιαστών Lagrange στην περίπτωση που έχουμε περισσότερους περιορισμούς ή συνθήκες.

Έστω $f, g_1, \dots, g_m : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^n$, έστω εσωτερικό σημείο $\mathbf{x}_0$ του A και έστω ότι οι $f, g_1, \dots, g_m$ έχουν συνεχείς μερικές παραγώγους πρώτης τάξης σε μια περιοχή του $\mathbf{x}_0$ και ότι τα διανύσματα $\nabla g_1(\mathbf{x}_0), \dots, \nabla g_m(\mathbf{x}_0)$ είναι γραμμικώς ανεξάρτητα. Θεωρούμε το σύνολο

$$S = \{\mathbf{x} \in A \mid g_1(\mathbf{x}) = c_1, \dots, g_m(\mathbf{x}) = c_m\},$$

δηλαδή την τομή ισοσταθμικών συνόλων των $g_1, \dots, g_m$. Υποθέτουμε ότι $\mathbf{x}_0 \in S$, δηλαδή ότι $g_1(\mathbf{x}_0) = c_1, \dots, g_m(\mathbf{x}_0) = c_m$. Υποθέτουμε, επίσης, ότι το $\mathbf{x}_0$ είναι σημείο τοπικού μεγίστου (ελαχίστου) της f περιορισμένης στο σύνολο S , δηλαδή ότι υπάρχει κάποιο $\delta > 0$ ώστε να ισχύει $f(\mathbf{x}) \leq (\geq) f(\mathbf{x}_0)$ για κάθε $\mathbf{x} \in S$ με $\|\mathbf{x} - \mathbf{x}_0\| < \delta$. Τότε υπάρχουν αριθμοί $\lambda_1, \dots, \lambda_m$ ώστε

$$\nabla f(\mathbf{x}_0) = \lambda_1 \nabla g_1(\mathbf{x}_0) + \dots + \lambda_m \nabla g_m(\mathbf{x}_0).$$

Οι αριθμοί $\lambda_1, \dots, \lambda_m$ ονομάζονται πολλαπλασιαστές Lagrange.

Παράδειγμα: Θέλουμε να βρούμε την μέγιστη τιμή και την ελάχιστη τιμή της $f(x, y) = x + y + z$ με τους περιορισμούς $x^2 + y^2 = 1$ και $2x + z = 1$. Με άλλα λόγια θέλουμε τις ακρότατες τιμές της f στο σύνολο

$$S = \{(x, y, z) \mid x^2 + y^2 = 1, 2x + z = 1\}.$$

Το S είναι η τομή του κυλίνδρου με καρτεσιανή εξίσωση $x^2 + y^2 = 1$ και του επιπέδου με καρτεσιανή εξίσωση $2x + z = 1$ και άρα είναι μια έλλειψη, δηλαδή ένα κλειστό και φραγμένο σύνολο. Τώρα, η f , η οποία είναι συνεχής στο S , έχει οπωσδήποτε μέγιστη τιμή και ελάχιστη τιμή στο S . Η f καθώς και οι συναρτήσεις $g_1(x, y, z) = x^2 + y^2$ και $g_2(x, y, z) = 2x + z$ έχουν συνεχείς μερικές παραγώγους πρώτης τάξης σε περιοχή κάθε σημείου του S (και μάλιστα σε ολόκληρο το $\mathbb{R}^3$). Είναι εύκολο να δει κάποιος ότι τα διανύσματα

$$\nabla g_1(x, y, z) = (2x, 2y, 0), \quad \nabla g_2(x, y, z) = (2, 0, 1)$$

είναι γραμμικώς ανεξάρτητα σε κάθε $(x, y, z) \in S$.

Βρίσκουμε τα σημεία (x, y, z) τοπικού ακροτάτου της f περιορισμένης στο S με την μέθοδο των πολλαπλασιαστών Lagrange. Αν το $(x, y, z) \in S$ είναι σημείο τοπικού ακροτάτου της f περιορισμένης στο S , τότε υπάρχουν αριθμοί λ_1, λ_2 ώστε

$$\nabla f(x, y, z) = \lambda_1 \nabla g_1(x, y, z) + \lambda_2 \nabla g_2(x, y, z)$$

ή, ισοδύναμα,

$$(1, 1, 1) = \lambda_1(2x, 2y, 0) + \lambda_2(2, 0, 1).$$

Συνδυάζοντας με τις $x^2 + y^2 = 1$ και $2x + z = 1$, βρίσκουμε ότι το σημείο τοπικού ακροτάτου (x, y, z) είναι λύση του συστήματος.

$$\begin{cases} 2\lambda_1 x + 2\lambda_2 = 1 \\ 2\lambda_1 y = 1 \\ \lambda_2 = 1 \\ x^2 + y^2 = 1 \\ 2x + z = 1 \end{cases}$$

Άρα $\lambda_2 = 1$ και έχουμε το μικρότερο σύστημα

$$\begin{cases} 2\lambda_1 x = -1 \\ 2\lambda_1 y = 1 \\ x^2 + y^2 = 1 \\ 2x + z = 1 \end{cases}$$

Προφανώς είναι $\lambda_1 \neq 0$, οπότε από τις δύο πρώτες εξισώσεις του συστήματος παίρνουμε

$$x = -\frac{1}{2\lambda_1}, \quad y = \frac{1}{2\lambda_1}.$$

Τότε η τρίτη εξίσωση γίνεται

$$\frac{1}{2\lambda_1^2} = 1$$

και άρα $\lambda_1 = -1/\sqrt{2}$ ή $\lambda_1 = 1/\sqrt{2}$. Άρα βρίσκουμε δύο αντίστοιχα ζεύγη (x, y) και μαζί με την τέταρτη εξίσωση $2x + z = 1$ προκύπτουν δύο πιθανά σημεία τοπικού ακροτάτου στο S :

$$(1/\sqrt{2}, -1/\sqrt{2}, 1 - \sqrt{2}), \quad (-1/\sqrt{2}, 1/\sqrt{2}, 1 + \sqrt{2}).$$

Στο πρώτο σημείο η f έχει τιμή $1 - \sqrt{2}$ και στο δεύτερο σημείο έχει τιμή $1 + \sqrt{2}$.

Άρα οι τιμές $1 - \sqrt{2}$ και $1 + \sqrt{2}$ είναι η ελάχιστη τιμή και η μέγιστη τιμή της f στο S .

12η εβδομάδα.

Θα δούμε την απόδειξη του θεωρήματος πάνω στο οποίο βασίζεται η μέθοδος των πολλαπλασιαστών Lagrange. Για απλότητα θα υποθέσουμε ότι $n = 3$. Το θεώρημα λέει τα εξής.

Έστω $f, g : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^3$, έστω εσωτερικό σημείο (x_0, y_0, z_0) του A , έστω ότι οι f, g έχουν συνεχείς μερικές παραγώγους πρώτης τάξης σε μια περιοχή του (x_0, y_0, z_0) και έστω ότι $\nabla g(x_0, y_0, z_0) \neq (0, 0, 0)$. Θεωρούμε το ισοσταθμικό σύνολο

$$S = \{(x, y, z) \in A \mid g(x, y, z) = c\}$$

και υποθέτουμε ότι $(x_0, y_0, z_0) \in S$, δηλαδή ότι $g(x_0, y_0, z_0) = c$. Υποθέτουμε ότι το (x_0, y_0, z_0) είναι σημείο τοπικού μεγίστου (ελαχίστου) της f περιορισμένης στο σύνολο S , δηλαδή ότι υπάρχει κάποιος $\delta > 0$ ώστε να ισχύει $f(x, y, z) \leq (\geq) f(x_0, y_0, z_0)$ για κάθε $(x, y, z) \in S$ τέτοιου ώστε $((x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2)^{1/2} < \delta$. Τότε υπάρχει αριθμός λ ώστε

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0).$$

Για την απόδειξη θα θυμηθούμε κάποια πράγματα που είδαμε στο τέλος της 8ης εβδομάδας. Εκεί είχαμε αναφέρει (και θα το αποδείξουμε λίγο αργότερα) ότι το μέρος του S το οποίο βρίσκεται κοντά στο (x_0, y_0, z_0) αποτελεί μία επιφάνεια η οποία έχει εφαπτόμενο επίπεδο στο σημείο (x_0, y_0, z_0) και, επίσης, ότι για κάθε διάνυσμα $\mathbf{v}$ το οποίο έχει αρχή το (x_0, y_0, z_0) και βρίσκεται πάνω στο εφαπτόμενο επίπεδο υπάρχει καμπύλη $\sigma : I \rightarrow \mathbb{R}^3$ έτσι ώστε να είναι

$$\sigma(t_0) = (x_0, y_0, z_0), \quad \sigma'(t_0) = \mathbf{v}.$$

για κάποιο εσωτερικό σημείο t_0 του διαστήματος I και έτσι ώστε η τροχιά της σ να περιέχεται στο ισοσταθμικό σύνολο S .

Θεωρώντας αυτήν την καμπύλη σ παίρνουμε την σύνθεση $h = f \circ \sigma : I \rightarrow \mathbb{R}$ και παρατηρούμε τα εξής. Λόγω συνέχειας της σ , υπάρχει κάποιος $\delta' > 0$ τέτοιο ώστε, αν $t \in I$ και $|t - t_0| < \delta'$, τότε $\|\sigma(t) - \sigma(t_0)\| < \delta$, δηλαδή το σημείο $\sigma(t)$ ανήκει στο ισοσταθμικό σύνολο S και απέχει από το σημείο $\sigma(t_0) = (x_0, y_0, z_0)$ απόσταση $< \delta$. Επομένως, αν $t \in I$ και $|t - t_0| < \delta'$, τότε $f(\sigma(t)) \leq (\geq) f(x_0, y_0, z_0) = f(\sigma(t_0))$ και άρα $h(t) \leq (\geq) h(t_0)$. Αυτό σημαίνει ότι το εσωτερικό σημείο t_0 του διαστήματος I είναι σημείο τοπικού μεγίστου (ελαχίστου) της συνάρτησης h . Από το Θεώρημα του Fermat συνεπάγεται $h'(t_0) = 0$ ή, ισοδύναμα, $\nabla f(\sigma(t_0)) \cdot \sigma'(t_0) = 0$, δηλαδή

$$\nabla f(x_0, y_0, z_0) \cdot \mathbf{v} = 0.$$

Βλέπουμε, λοιπόν, ότι το διάνυσμα $\nabla f(x_0, y_0, z_0)$ είναι κάθετο σε κάθε διάνυσμα $\mathbf{v}$ το οποίο έχει αρχή το (x_0, y_0, z_0) και βρίσκεται πάνω στο εφαπτόμενο επίπεδο της ισοσταθμικής επιφάνειας S . Άρα το $\nabla f(x_0, y_0, z_0)$ είναι κάθετο στην S στο σημείο (x_0, y_0, z_0) . Επειδή, όμως, και το διάνυσμα $\nabla g(x_0, y_0, z_0)$ είναι κάθετο στην S στο ίδιο σημείο και είναι μη-μηδενικό διάνυσμα, καταλήγουμε στο ότι το $\nabla f(x_0, y_0, z_0)$ είναι πολλαπλάσιο του $\nabla g(x_0, y_0, z_0)$.

Επόμενο θέμα. Είναι προφανές ότι ο μοναδιαίος κύκλος στο xy -επίπεδο, δηλαδή το σύνολο των σημείων (x, y) που ικανοποιούν την $x^2 + y^2 = 1$, δεν μπορεί να είναι γράφημα κάποιας συνάρτησης $y = f(x)$, όπου $f : [-1, 1] \rightarrow \mathbb{R}$, ούτε γράφημα κάποιας συνάρτησης $x = g(y)$, όπου

$g : [-1, 1] \rightarrow \mathbb{R}$. Διότι υπάρχουν κατακόρυφες ευθείες και οριζόντιες ευθείες οι οποίες τέμνουν τον κύκλο σε δύο σημεία η καθεμιά. Αυτό, όμως, το “πρόβλημα” δεν υφίσταται όταν πάρουμε κατάλληλα μικρά κομμάτια του κύκλου.

Για παράδειγμα, αν (x_0, y_0) είναι κάποιο σημείο του κύκλου διαφορετικό από τα σημεία $(-1, 0)$ και $(1, 0)$, μπορούμε να βρούμε εύκολα κάποιο ορθογώνιο παραλληλόγραμμο $U \times V$, όπου U είναι ένα διάστημα με το x_0 ως εσωτερικό του σημείο και V είναι κάποιο διάστημα με το y_0 ως εσωτερικό του σημείο, έτσι ώστε το μέρος του κύκλου που βρίσκεται στο ορθογώνιο $U \times V$ να είναι γράφημα μιας συνάρτησης $y = f(x)$, όπου $f : U \rightarrow V$. Προφανώς, ο τύπος της f πρέπει να είναι $y = f(x) = \sqrt{1 - x^2}$ για κάθε $x \in U$ ή $y = f(x) = -\sqrt{1 - x^2}$ για κάθε $x \in U$ (ανάλογα με το αν το σημείο (x_0, y_0) βρίσκεται στο άνω ημικύκλιο ή στο κάτω ημικύκλιο). Μπορούμε, λοιπόν, να λύσουμε την εξίσωση $x^2 + y^2 = 1$ ως προς το $y \in V$ συναρτήσει του $x \in U$: για κάθε $x \in U$ υπάρχει ένα και μοναδικό $y \in V$ ώστε να ισχύει $x^2 + y^2 = 1$. Παρατηρούμε ότι, αν θεωρήσουμε την συνάρτηση $F(x, y) = x^2 + y^2$, τα σημεία (x_0, y_0) στα οποία αναφερόμαστε είναι εκείνα για τα οποία το διάνυσμα $\nabla F(x_0, y_0) = (2x_0, 2y_0)$, το οποίο είναι κάθετο στον κύκλο στο σημείο (x_0, y_0) , δεν είναι οριζόντιο, δηλαδή δεν έχει δεύτερη συντεταγμένη ίση με 0.

Ομοίως, αν (x_0, y_0) είναι κάποιο σημείο του κύκλου διαφορετικό από τα σημεία $(0, -1)$ και $(0, 1)$, μπορούμε να βρούμε κάποιο ορθογώνιο παραλληλόγραμμο $U \times V$, όπου U είναι ένα διάστημα με το x_0 ως εσωτερικό του σημείο και V είναι κάποιο διάστημα με το y_0 ως εσωτερικό του σημείο, έτσι ώστε το μέρος του κύκλου που βρίσκεται στο ορθογώνιο $U \times V$ να είναι γράφημα μιας συνάρτησης $x = g(y)$, όπου $g : V \rightarrow U$. Προφανώς, ο τύπος της g πρέπει να είναι $x = g(y) = \sqrt{1 - y^2}$ για κάθε $y \in V$ ή $x = g(y) = -\sqrt{1 - y^2}$ για κάθε $y \in V$ (ανάλογα με το αν το σημείο (x_0, y_0) βρίσκεται στο δεξιό ημικύκλιο ή στο αριστερό ημικύκλιο). Μπορούμε, λοιπόν, να λύσουμε την εξίσωση $x^2 + y^2 = 1$ ως προς το $x \in U$ συναρτήσει του $y \in V$: για κάθε $y \in V$ υπάρχει ένα και μοναδικό $x \in U$ ώστε να ισχύει $x^2 + y^2 = 1$. Όπως στην προηγούμενη περίπτωση, παρατηρούμε ότι τα σημεία (x_0, y_0) στα οποία αναφερόμαστε είναι εκείνα για τα οποία το διάνυσμα $\nabla F(x_0, y_0) = (2x_0, 2y_0)$ δεν είναι κατακόρυφο, δηλαδή δεν έχει πρώτη συντεταγμένη ίση με 0.

Τα προηγούμενα ισχύουν γενικότερα. Αν έχουμε μια καμπύλη με κάποιες ελάχιστες προϋποθέσεις, τότε “τοπικά” μπορούμε να θεωρήσουμε ότι η καμπύλη είναι το γράφημα μιας συνάρτησης του x ή το γράφημα μιας συνάρτησης του y . Αυτό είναι το περιεχόμενο του εξής θεωρήματος.

Έστω $F : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^2$, έστω εσωτερικό σημείο (x_0, y_0) του A και έστω

$$S = \{(x, y) \in A \mid F(x, y) = c\}$$

ένα ισοσταθμικό σύνολο της F . Υποθέτουμε ότι το (x_0, y_0) ανήκει στο S , δηλαδή

$$F(x_0, y_0) = c.$$

Επίσης, υποθέτουμε ότι η F έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε μια περιοχή του (x_0, y_0) και ότι

$$\nabla F(x_0, y_0) \neq (0, 0).$$

Δηλαδή, $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$ ή $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$.

(i) Έστω $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$. Τότε υπάρχει διάστημα U με το x_0 ως εσωτερικό του σημείο και διάστημα V με το y_0 ως εσωτερικό του σημείο ώστε για κάθε $x \in U$ να υπάρχει ένα και μοναδικό $y \in V$ ώστε το σημείο (x, y) να ανήκει στο S ή, ισοδύναμα, να ισχύει $F(x, y) = c$. Με άλλα λόγια, το μέρος του συνόλου S που περιέχεται στο ορθογώνιο $U \times V$ είναι γράφημα μιας συνάρτησης $f : U \rightarrow V$, όπου για κάθε $x \in U$ το $f(x)$ είναι το μοναδικό $y \in V$ για το οποίο ισχύει $F(x, y) = c$.

(ii) Έστω $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$. Τότε υπάρχει διάστημα U με το x_0 ως εσωτερικό του σημείο και διάστημα V με το y_0 ως εσωτερικό του σημείο ώστε για κάθε $y \in V$ να υπάρχει ένα και μοναδικό $x \in U$ ώστε το σημείο (x, y) να ανήκει στο S ή, ισοδύναμα, να ισχύει $F(x, y) = c$. Με άλλα λόγια, το μέρος του συνόλου S που περιέχεται στο ορθογώνιο $U \times V$ είναι γράφημα μιας συνάρτησης $g : V \rightarrow U$, όπου για κάθε $y \in V$ το $g(y)$ είναι το μοναδικό $x \in U$ για το οποίο ισχύει $F(x, y) = c$.

Στην περίπτωση (i), όταν $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$, λέμε ότι “τοπικά”, δηλαδή σε μια περιοχή του (x_0, y_0) , μπορούμε να λύσουμε την εξίσωση $F(x, y) = c$ και να εκφράσουμε το y ως συνάρτηση του x . Αναλόγως, στην περίπτωση (ii), όταν $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$, λέμε ότι “τοπικά” μπορούμε να λύσουμε την εξίσωση $F(x, y) = c$ και να εκφράσουμε το x ως συνάρτηση του y .

Το θεώρημα έχει και συνέχεια.

Στην περίπτωση (i) η συνάρτηση $f : U \rightarrow V$ έχει συνεχή παράγωγο ως προς το x στο διάστημα U η οποία δίνεται από τον τύπο

$$\frac{df}{dx}(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))} \quad \text{για κάθε } x \in U.$$

Στην περίπτωση (ii) η συνάρτηση $g : V \rightarrow U$ έχει συνεχή παράγωγο ως προς το y στο διάστημα V η οποία δίνεται από τον τύπο

$$\frac{dg}{dy}(y) = -\frac{\frac{\partial F}{\partial y}(g(y), y)}{\frac{\partial F}{\partial x}(g(y), y)} \quad \text{για κάθε } y \in V.$$

Παράδειγμα. Θεωρούμε την εξίσωση

$$x^2 + xy + e^y = 2$$

και θέλουμε να δούμε αν μπορούμε να την λύσουμε ως προς το y ως συνάρτηση του x κοντά στο σημείο $(1, 0)$. Κατ’ αρχάς το σημείο $(1, 0)$ ικανοποιεί την εξίσωση, αλλιώς το πρόβλημα δεν θα είχε νόημα. Σύμφωνα με το θεώρημα πρέπει να δούμε αν ικανοποιούνται κάποιες προϋποθέσεις. Η συνάρτηση

$$F(x, y) = x^2 + xy + e^y$$

έχει συνεχείς μερικές παραγώγους σε κάθε σημείο του $\mathbb{R}^2$ (και άρα και σε οποιαδήποτε περιοχή του σημείου $(1, 0)$):

$$\frac{\partial F}{\partial x} = 2x + y, \quad \frac{\partial F}{\partial y} = x + e^y.$$

Επειδή θέλουμε να λύσουμε ως προς y ελέγχουμε αν η μερική παράγωγος ως προς το y στο σημείο $(1, 0)$ είναι $\neq 0$:

$$\frac{\partial F}{\partial y}(1, 0) = 2 \neq 0.$$

Άρα το θεώρημα λέει ότι υπάρχει διάστημα U με το 1 ως εσωτερικό του σημείο και διάστημα V με το 0 ως εσωτερικό του σημείο ώστε για κάθε $x \in U$ να υπάρχει ένα και μοναδικό $y \in V$ ώστε να ισχύει $F(x, y) = 2$. Δηλαδή ορίζεται συνάρτηση $f : U \rightarrow V$, όπου για κάθε $x \in U$ το $f(x)$ είναι το μοναδικό $y \in V$ για το οποίο ισχύει $F(x, y) = 2$. Έχουμε λοιπόν

$$x^2 + xf(x) + e^{f(x)} = F(x, f(x)) = 2 \quad \text{για κάθε } x \in U.$$

Η συνάρτηση f έχει συνεχή παράγωγο ως προς x στην περιοχή U και ο τύπος της είναι

$$\frac{df}{dx}(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))} = -\frac{2x + f(x)}{x + e^{f(x)}} \quad \text{για κάθε } x \in U.$$

Το θεώρημα αυτό ονομάζεται **Θεώρημα Πεπλεγμένης Συνάρτησης** και έχει μια γενικότερη μορφή, όταν η συνάρτηση F έχει $n + 1$ μεταβλητές,

$$F(x_1, \dots, x_n, z),$$

και θέλουμε να λύσουμε την εξίσωση

$$F(x_1, \dots, x_n, z) = c$$

ως προς την τελευταία μεταβλητή z ως συνάρτηση των υπόλοιπων μεταβλητών $(x_1, \dots, x_n)$.
Εστω $F : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^{n+1}$, έστω εσωτερικό σημείο $(x_{01}, \dots, x_{0n}, z_0)$ του A και έστω

$$S = \{(x_1, \dots, x_n, z) \in A \mid F(x_1, \dots, x_n, z) = c\}$$

ένα ισοσταθμικό σύνολο της F . Υποθέτουμε ότι το $(x_{01}, \dots, x_{0n}, z_0)$ ανήκει στο S , δηλαδή

$$F(x_{01}, \dots, x_{0n}, z_0) = c.$$

Επίσης, υποθέτουμε ότι η F έχει συνεχείς μερικές παραγώγους πρώτης τάξης σε μια περιοχή του $(x_{01}, \dots, x_{0n}, z_0)$ και ότι

$$\frac{\partial F}{\partial z}(x_{01}, \dots, x_{0n}, z_0) \neq 0.$$

Τότε υπάρχει περιοχή U στον $\mathbb{R}^n$ με το σημείο $(x_{01}, \dots, x_{0n})$ ως εσωτερικό της σημείο και διάστημα V στο $\mathbb{R}$ με το z_0 ως εσωτερικό του σημείο ώστε για κάθε $(x_1, \dots, x_n) \in U$ να υπάρχει ένα και μοναδικό $z \in V$ ώστε το σημείο $(x_1, \dots, x_n, z)$ να ανήκει στο S ή, ισοδύναμα, να ισχύει $F(x_1, \dots, x_n, z) = c$. Με άλλα λόγια, το μέρος του συνόλου S που περιέχεται στο καρτεσιανό γινόμενο $U \times V$ είναι γράφημα μιας συνάρτησης $h : U \rightarrow V$, όπου για κάθε $(x_1, \dots, x_n) \in U$ το $h(x_1, \dots, x_n)$ είναι το μοναδικό $z \in V$ για το οποίο ισχύει $F(x_1, \dots, x_n, z) = c$ και επομένως έχουμε ότι

$$F(x_1, \dots, x_n, h(x_1, \dots, x_n)) = c \quad \text{για κάθε } (x_1, \dots, x_n) \in U.$$

Η συνάρτηση $h : U \rightarrow V$ έχει συνεχείς μερικές παραγώγους πρώτης τάξης ως προς τις μεταβλητές $x_1, \dots, x_n$ στην περιοχή U οι οποίες δίνονται από τους τύπους

$$\frac{\partial h}{\partial x_j}(x_1, \dots, x_n) = -\frac{\frac{\partial F}{\partial x_j}(x_1, \dots, x_n, h(x_1, \dots, x_n))}{\frac{\partial F}{\partial z}(x_1, \dots, x_n, h(x_1, \dots, x_n))} \quad \text{για κάθε } (x_1, \dots, x_n) \in U.$$

Παράδειγμα. Θεωρούμε την εξίσωση

$$xy + z^2 + 3xz^5 = 4$$

και θέλουμε να την λύσουμε ως προς το z ως συνάρτηση των x, y κοντά στο σημείο $(1, 0, 1)$. Το σημείο $(1, 0, 1)$ ικανοποιεί την εξίσωση και ελέγχουμε τις προϋποθέσεις του θεωρήματος. Η συνάρτηση

$$F(x, y) = xy + z^2 + 3xz^5$$

έχει συνεχείς μερικές παραγώγους σε κάθε σημείο του $\mathbb{R}^3$ (και άρα και σε οποιαδήποτε περιοχή του σημείου $(1, 0, 1)$):

$$\frac{\partial F}{\partial x} = y + 3z^5, \quad \frac{\partial F}{\partial y} = x, \quad \frac{\partial F}{\partial z} = 2z + 15xz^4.$$

Επειδή θέλουμε να λύσουμε ως προς z ελέγχουμε αν η μερική παράγωγος ως προς το z στο σημείο $(1, 0, 1)$ είναι $\neq 0$:

$$\frac{\partial F}{\partial z}(1, 0, 1) = 17 \neq 0.$$

Άρα υπάρχει περιοχή U στον $\mathbb{R}^2$ (π.χ. δίσκος ή ορθογώνιο) με το $(1, 0)$ ως εσωτερικό της σημείο και διάστημα V με το 1 ως εσωτερικό του σημείο ώστε για κάθε $(x, y) \in U$ να υπάρχει ένα και μοναδικό $z \in V$ ώστε να ισχύει $F(x, y, z) = 4$. Δηλαδή ορίζεται συνάρτηση $h : U \rightarrow V$, όπου για κάθε $(x, y) \in U$ το $h(x, y)$ είναι το μοναδικό $z \in V$ για το οποίο ισχύει $F(x, y, z) = 4$. Έχουμε λοιπόν

$$xy + h(x, y)^2 + 3xh(x, y)^5 = F(x, y, h(x, y)) = 4 \quad \text{για κάθε } (x, y) \in U.$$

Η συνάρτηση h έχει συνεχείς μερικές παραγώγους πρώτης τάξης ως προς x και y στην περιοχή U και οι τύποι τους είναι

$$\frac{\partial h}{\partial x}(x, y) = -\frac{\frac{\partial F}{\partial x}(x, y, h(x, y))}{\frac{\partial F}{\partial z}(x, y, h(x, y))} = -\frac{y + 3h(x, y)^5}{2h(x, y) + 15xh(x, y)^4},$$

$$\frac{\partial h}{\partial y}(x, y) = -\frac{\frac{\partial F}{\partial y}(x, y, h(x, y))}{\frac{\partial F}{\partial z}(x, y, h(x, y))} = -\frac{x}{2h(x, y) + 15xh(x, y)^4}$$

για κάθε $(x, y) \in U$.

13η εβδομάδα.

Για να καταλάβουμε καλύτερα τον ρόλο της υπόθεσης $\frac{\partial F}{\partial z} \neq 0$ στο Θεώρημα Πεπλεγμένης Συνάρτησης, ας δούμε την απλή ειδική περίπτωση γραμμικής συνάρτησης

$$F(x_1, \dots, x_n, z) = a_1 x_1 + \dots + a_n x_n + bz.$$

Για να μπορούμε να λύσουμε την εξίσωση

$$a_1 x_1 + \dots + a_n x_n + bz = c$$

ως προς το z ως συνάρτηση των $x_1, \dots, x_n$ χρειάζεται να υποθέσουμε ότι $b \neq 0$ ώστε να πάρουμε

$$z = -\frac{a_1}{b} x_1 - \dots - \frac{a_n}{b} x_n + \frac{c}{b}.$$

Παρατηρήστε τώρα ότι

$$\frac{\partial F}{\partial z} = b,$$

οπότε η συνθήκη $b \neq 0$ ισοδυναμεί με την $\frac{\partial F}{\partial z} \neq 0$.

Εν όψει της γενίκευσης του Θεωρήματος Πεπλεγμένης Συνάρτησης σε περισσότερες από μία εξισώσεις, ας δούμε την ειδική περίπτωση m γραμμικών εξισώσεων:

$$a_{11}x_1 + \dots + a_{1n}x_n + b_{11}z_1 + \dots + b_{1m}z_m = c_1$$

$$\vdots$$

$$a_{m1}x_1 + \dots + a_{mn}x_n + b_{m1}z_1 + \dots + b_{mm}z_m = c_m$$

Έχουμε ένα γραμμικό σύστημα m εξισώσεων με m αγνώστους: τα $z_1, \dots, z_m$. Αν μεταφέρουμε τους όρους που περιέχουν τα $x_1, \dots, x_n$ στην δεξιά μεριά, τότε παίρνουμε ένα μη-ομογενές σύστημα το οποίο για να έχει λύση χρειάζεται να υποθέσουμε ότι

$$\det \begin{bmatrix} b_{11} & \dots & b_{1m} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mm} \end{bmatrix} \neq 0.$$

Αν, λοιπόν, ονομάσουμε $F_1(x_1, \dots, x_n, z_1, \dots, z_m), \dots, F_m(x_1, \dots, x_n, z_1, \dots, z_m)$ τις παραστάσεις στην αριστερή μεριά των εξισώσεων, βλέπουμε ότι η συνθήκη γράφεται

$$\det \begin{bmatrix} \frac{\partial F_1}{\partial z_1} & \dots & \frac{\partial F_1}{\partial z_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial z_1} & \dots & \frac{\partial F_m}{\partial z_m} \end{bmatrix} \neq 0.$$

Μετά από αυτά τα εισαγωγικά, ορίστε το **γενικό Θεώρημα Πεπλεγμένης Συνάρτησης**.

Θεωρούμε συναρτήσεις $F_1, \dots, F_m : A \rightarrow \mathbb{R}$, όπου $A \subseteq \mathbb{R}^{n+m}$ καθώς και το εσωτερικό σημείο

$(\mathbf{x}_0, \mathbf{z}_0) = (x_{01}, \dots, x_{0n}, z_{01}, \dots, z_{0m})$ του A . Θεωρούμε και το σύνολο S των σημείων $(\mathbf{x}, \mathbf{z}) = (x_1, \dots, x_n, z_1, \dots, z_m) \in A$ τα οποία ικανοποιούν τις m εξισώσεις

$$\begin{aligned} F_1(\mathbf{x}, \mathbf{z}) &= F_1(x_1, \dots, x_n, z_1, \dots, z_m) = c_1 \\ &\vdots \\ F_m(\mathbf{x}, \mathbf{z}) &= F_m(x_1, \dots, x_n, z_1, \dots, z_m) = c_m \end{aligned}$$

Υποθέτουμε ότι το σημείο $(\mathbf{x}_0, \mathbf{z}_0) = (x_{01}, \dots, x_{0n}, z_{01}, \dots, z_{0m})$ ανήκει στο S , δηλαδή ικανοποιεί όλες τις παραπάνω εξισώσεις.

Επίσης, υποθέτουμε ότι οι $F_1, \dots, F_m$ έχουν συνεχείς μερικές παραγώγους πρώτης τάξης ως προς όλες τις μεταβλητές σε μια περιοχή του σημείου $(\mathbf{x}_0, \mathbf{z}_0) = (x_{01}, \dots, x_{0n}, z_{01}, \dots, z_{0m})$ και ότι στο σημείο αυτό ισχύει

$$\det \begin{bmatrix} \frac{\partial F_1}{\partial z_1}(\mathbf{x}_0, \mathbf{z}_0) & \dots & \frac{\partial F_1}{\partial z_m}(\mathbf{x}_0, \mathbf{z}_0) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial z_1}(\mathbf{x}_0, \mathbf{z}_0) & \dots & \frac{\partial F_m}{\partial z_m}(\mathbf{x}_0, \mathbf{z}_0) \end{bmatrix} \neq 0. \quad (72)$$

Τότε υπάρχει περιοχή U στον $\mathbb{R}^n$ με το $\mathbf{x}_0 = (x_{01}, \dots, x_{0n})$ ως εσωτερικό της σημείο και περιοχή V στον $\mathbb{R}^m$ με το $\mathbf{z}_0 = (z_{01}, \dots, z_{0m})$ ως εσωτερικό της σημείο ώστε για κάθε $\mathbf{x} = (x_1, \dots, x_n) \in U$ να υπάρχει ακριβώς ένα $\mathbf{z} = (z_1, \dots, z_m) \in V$ ώστε το σημείο $(\mathbf{x}, \mathbf{z}) = (x_1, \dots, x_n, z_1, \dots, z_m)$ να ανήκει στο S ή, ισοδύναμα, να ικανοποιεί το σύστημα των εξισώσεων. Με άλλα λόγια, το μέρος του συνόλου S που περιέχεται στο καρτεσιανό γινόμενο $U \times V$ είναι το γράφημα μιας συνάρτησης $\mathbf{z} = (z_1, \dots, z_m) : U \rightarrow V$, όπου για κάθε $\mathbf{x} = (x_1, \dots, x_n) \in U$ το

$$\mathbf{z}(\mathbf{x}) = (z_1(x_1, \dots, x_n), \dots, z_m(x_1, \dots, x_n))$$

είναι το μοναδικό $\mathbf{z} = (z_1, \dots, z_m) \in V$ το οποίο ικανοποιεί το σύστημα των εξισώσεων και επομένως έχουμε ότι

$$\begin{aligned} F_1(\mathbf{x}, \mathbf{z}(\mathbf{x})) &= F_1(x_1, \dots, x_n, z_1(x_1, \dots, x_n), \dots, z_m(x_1, \dots, x_n)) = c_1 \\ &\vdots \\ F_m(\mathbf{x}, \mathbf{z}(\mathbf{x})) &= F_m(x_1, \dots, x_n, z_1(x_1, \dots, x_n), \dots, z_m(x_1, \dots, x_n)) = c_m \end{aligned}$$

για κάθε $\mathbf{x} = (x_1, \dots, x_n) \in U$.

Η συνάρτηση $\mathbf{z} = (z_1, \dots, z_m) : U \rightarrow V$ έχει συνεχείς μερικές παραγώγους πρώτης τάξης ως προς τις μεταβλητές $x_1, \dots, x_n$ στην περιοχή U οι οποίες δίνονται από τους τύπους

$$\frac{\partial z_i}{\partial x_j}(\mathbf{x}) = - \frac{\det \begin{bmatrix} \frac{\partial F_1}{\partial z_1}(\mathbf{x}, \mathbf{z}(\mathbf{x})) & \dots & \frac{\partial F_1}{\partial x_j}(\mathbf{x}, \mathbf{z}(\mathbf{x})) & \dots & \frac{\partial F_1}{\partial z_m}(\mathbf{x}, \mathbf{z}(\mathbf{x})) \\ \vdots & & \vdots & & \vdots \\ \frac{\partial F_m}{\partial z_1}(\mathbf{x}, \mathbf{z}(\mathbf{x})) & \dots & \frac{\partial F_m}{\partial x_j}(\mathbf{x}, \mathbf{z}(\mathbf{x})) & \dots & \frac{\partial F_m}{\partial z_m}(\mathbf{x}, \mathbf{z}(\mathbf{x})) \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F_1}{\partial z_1}(\mathbf{x}, \mathbf{z}(\mathbf{x})) & \dots & \frac{\partial F_1}{\partial z_m}(\mathbf{x}, \mathbf{z}(\mathbf{x})) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial z_1}(\mathbf{x}, \mathbf{z}(\mathbf{x})) & \dots & \frac{\partial F_m}{\partial z_m}(\mathbf{x}, \mathbf{z}(\mathbf{x})) \end{bmatrix}}$$

για κάθε $\mathbf{x} = (x_1, \dots, x_n) \in U$, όπου ο πίνακας στον αριθμητή διαφέρει από τον πίνακα στον παρονομαστή μόνο ως προς την i στήλη (οι μερικές παράγωγοι ως προς z_i έχουν αντικατασταθεί από τις μερικές παραγώγους ως προς x_j).

Παράδειγμα: Θεωρούμε το σύστημα των εξισώσεων

$$\begin{aligned} x + y + z^3 w &= 2 \\ xyz + w^2 &= 4 \end{aligned}$$

και θέλουμε να δούμε αν λύνεται ως προς τις μεταβλητές z, w ως συναρτήσεις των μεταβλητών x, y κοντά στο σημείο $(x_0, y_0, z_0, w_0) = (0, 0, 1, 2)$.

Θεωρούμε τις συναρτήσεις

$$F_1(x, y, z, w) = x + y + z^3 w, \quad F_2(x, y, z, w) = xyz + w^2$$

οι οποίες έχουν συνεχείς μερικές παραγώγους πρώτης τάξης

$$\begin{aligned} \frac{\partial F_1}{\partial x} &= 1, & \frac{\partial F_1}{\partial y} &= 1, & \frac{\partial F_1}{\partial z} &= 3z^2 w, & \frac{\partial F_1}{\partial w} &= z^3, \\ \frac{\partial F_2}{\partial x} &= yz, & \frac{\partial F_2}{\partial y} &= xz, & \frac{\partial F_2}{\partial z} &= xy, & \frac{\partial F_2}{\partial w} &= 2w \end{aligned}$$

σε ολόκληρο τον $\mathbb{R}^{2+2} = \mathbb{R}^4$. Επίσης, το σημείο $(x_0, y_0, z_0, w_0) = (0, 0, 1, 2)$ ικανοποιεί τις εξισώσεις. Ελέγχουμε την συνθήκη με την μη-μηδενική ορίζουσα στο σημείο $(x_0, y_0, z_0, w_0) = (0, 0, 1, 2)$:

$$\det \begin{bmatrix} \frac{\partial F_1}{\partial z} & \frac{\partial F_1}{\partial w} \\ \frac{\partial F_2}{\partial z} & \frac{\partial F_2}{\partial w} \end{bmatrix} = \det \begin{bmatrix} 3z^2 w & z^3 \\ xy & 2w \end{bmatrix} = 6z^2 w^2 - xyz^3$$

και άρα

$$\det \begin{bmatrix} \frac{\partial F_1}{\partial z}(0, 0, 1, 2) & \frac{\partial F_1}{\partial w}(0, 0, 1, 2) \\ \frac{\partial F_2}{\partial z}(0, 0, 1, 2) & \frac{\partial F_2}{\partial w}(0, 0, 1, 2) \end{bmatrix} = 24 \neq 0.$$

Άρα υπάρχει περιοχή U στον $\mathbb{R}^2$ με το $(x_0, y_0) = (0, 0)$ ως εσωτερικό της σημείο και περιοχή V στον $\mathbb{R}^2$ με το $(z_0, w_0) = (1, 2)$ ως εσωτερικό της σημείο ώστε για κάθε $(x, y) \in U$ να υπάρχει ακριβώς ένα $(z, w) \in V$ ώστε το σημείο (x, y, z, w) να ικανοποιεί το σύστημα των εξισώσεων. Με άλλα λόγια, ορίζεται συνάρτηση $(z, w) : U \rightarrow V$, όπου για κάθε $(x, y) \in U$ το

$$(z(x, y), w(x, y))$$

είναι το μοναδικό $(z, w) \in V$ το οποίο ικανοποιεί το σύστημα των εξισώσεων και επομένως έχουμε ότι

$$\begin{aligned} F_1(x, y, z(x, y), w(x, y)) &= x + y + z(x, y)^3 w(x, y) = 2 \\ F_2(x, y, z(x, y), w(x, y)) &= xyz(x, y) + w(x, y)^2 = 4 \end{aligned} \tag{73}$$

για κάθε $(x, y) \in U$.

Η συνάρτηση $(z, w) : U \rightarrow V$ έχει συνεχείς μερικές παραγώγους πρώτης τάξης ως προς τις μεταβλητές x, y στην περιοχή U . Αυτό φυσικά σημαίνει ότι οι συντεταγμένες συναρτήσεις z, w έχουν συνεχείς μερικές παραγώγους ως προς τις x, y . Για να τις βρούμε μπορούμε να χρησιμοποιήσουμε τους γενικούς τύπους του Θεωρήματος Πεπλεγμένης Συνάρτησης αλλά προτιμάμε, αντί να αποστηθίσουμε γενικούς τύπους, να ακολουθήσουμε την διαδικασία με την οποία προκύπτουν και οι γενικοί τύποι. Παραγωγίζουμε τις εξισώσεις (73) ως προς x, y . Παραγωγίζοντας ως προς x βρίσκουμε (ως συνήθως, για να εξοικονομήσουμε χώρο γράφουμε z, w αντί $z(x, y), w(x, y)$)

$$\begin{aligned} 1 + 3z^2 w \frac{\partial z}{\partial x} + z^3 \frac{\partial w}{\partial x} &= 0 \\ yz + xy \frac{\partial z}{\partial x} + 2w \frac{\partial w}{\partial x} &= 0 \end{aligned}$$

Αυτό είναι ένα μη-ομογενές γραμμικό σύστημα με “αγνώστους” $\frac{\partial z}{\partial x}, \frac{\partial w}{\partial x}$ και, βάσει των τύπων του Cramer, βρίσκουμε

$$\frac{\partial z}{\partial x} = \frac{\begin{bmatrix} -1 & z^3 \\ -yz & 2w \end{bmatrix}}{\begin{bmatrix} 3z^2 w & z^3 \\ xy & 2w \end{bmatrix}} = \frac{-2w + yz^4}{6z^2 w^2 - xyz^3}, \quad \frac{\partial w}{\partial x} = \frac{\begin{bmatrix} 3z^2 w & -1 \\ xy & -yz \end{bmatrix}}{\begin{bmatrix} 3z^2 w & z^3 \\ xy & 2w \end{bmatrix}} = \frac{-3yz^3 w + xy}{6z^2 w^2 - xyz^3}.$$

Φυσικά, στην θέση των z, w κατανοούμε ότι βρίσκονται οι συναρτήσεις $z(x, y), w(x, y)$. Ομοίως, παραγωγίζουμε τις (73) ως προς y και έχουμε

$$\begin{aligned} 1 + 3z^2w \frac{\partial z}{\partial y} + z^3 \frac{\partial w}{\partial y} &= 0 \\ xz + xy \frac{\partial z}{\partial y} + 2w \frac{\partial w}{\partial y} &= 0 \end{aligned}$$

Άρα

$$\frac{\partial z}{\partial y} = \frac{\begin{bmatrix} -1 & z^3 \\ -xz & 2w \end{bmatrix}}{\begin{bmatrix} 3z^2w & z^3 \\ xy & 2w \end{bmatrix}} = \frac{-2w + xz^4}{6z^2w^2 - xyz^3}, \quad \frac{\partial w}{\partial y} = \frac{\begin{bmatrix} 3z^2w & -1 \\ xy & -xz \end{bmatrix}}{\begin{bmatrix} 3z^2w & z^3 \\ xy & 2w \end{bmatrix}} = \frac{-3xz^3w + xy}{6z^2w^2 - xyz^3}.$$

Στο σημείο αυτό ας θυμηθούμε μια παλιότερη ορολογία στο μάθημα αυτό. Αν έχουμε διανυσματική συνάρτηση $\mathbf{F} = (F_1, \dots, F_m)$ και καθεμία από τις συντεταγμένες πραγματικές συναρτήσεις $F_1, \dots, F_m$ είναι συνάρτηση n μεταβλητών $x_1, \dots, x_n$ τότε ορίζεται ο $m \times n$ πίνακας

$$D\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial F_1}{\partial x_n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial F_m}{\partial x_n}(\mathbf{x}) \end{bmatrix}$$

και ονομάζεται *παράγωγος της $\mathbf{F}$ στο $\mathbf{x}$* .

Στην περίπτωση $m = n$, δηλαδή όταν ο αριθμός των συντεταγμένων συναρτήσεων $F_1, \dots, F_n$ είναι ίδιος με τον αριθμό των μεταβλητών $x_1, \dots, x_n$, τότε ο πίνακας είναι τετραγωνικός $n \times n$ και χρησιμοποιούμε δύο (ισοδύναμα) σύμβολα για την ορίζουσα του πίνακα $D\mathbf{F}(\mathbf{x})$:

$$J\mathbf{F}(\mathbf{x}) = \frac{\partial(F_1, \dots, F_n)}{\partial(x_1, \dots, x_n)}(\mathbf{x}) = \det D\mathbf{F}(\mathbf{x}) = \det \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial F_1}{\partial x_n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial F_n}{\partial x_n}(\mathbf{x}) \end{bmatrix}$$

Η ορίζουσα αυτή, είτε με το σύμβολο $J\mathbf{F}(\mathbf{x})$ είτε με το $\frac{\partial(F_1, \dots, F_n)}{\partial(x_1, \dots, x_n)}(\mathbf{x})$ ονομάζεται **Ιακωβιανή ορίζουσα της $\mathbf{F}$ στο $\mathbf{x}$** .

Σύμφωνα με την τελευταία ορολογία, η σχέση (72) του Θεωρήματος Πεπλεγμένης Συνάρτησης γράφεται και ως εξής

$$J_{\mathbf{z}}\mathbf{F}(\mathbf{x}_0, \mathbf{z}_0) \neq 0 \quad \text{και} \quad \frac{\partial(F_1, \dots, F_m)}{\partial(z_1, \dots, z_m)}(\mathbf{x}_0, \mathbf{z}_0) \neq 0.$$

Στο σύμβολο $J_{\mathbf{z}}\mathbf{F}$ έχουμε βάλει το $\mathbf{z}$ ως δείκτη για να φαίνεται ότι παραγωγίζουμε τις $F_1, \dots, F_m$ ως προς τις μεταβλητές $z_1, \dots, z_m$ κρατώντας τις υπόλοιπες μεταβλητές $x_1, \dots, x_n$ σταθερές.

Το τελευταίο θεώρημα του μαθήματος έχει να κάνει με την αντιστρεψιμότητα μιας συνάρτησης. Ας δούμε τί γίνεται στην περίπτωση πραγματικής συνάρτησης μίας μεταβλητής. Έστω ότι η $f : I \rightarrow \mathbb{R}$ είναι *συνεχώς παραγωγίσιμη* σε κάποιο διάστημα I και έστω ότι σε κάποιο εσωτερικό σημείο x_0 του I ισχύει $f'(x_0) > 0$. Επειδή η f' είναι συνεχής, υπάρχει κάποιο διάστημα U με το x_0 εσωτερικό του σημείο ώστε να ισχύει $f'(x) > 0$ για κάθε $x \in U$. Άρα η f είναι γνησίως αύξουσα (και συνεχής) στο διάστημα U , οπότε απεικονίζει το U σε κάποιο διάστημα V με το $y_0 = f(x_0)$ εσωτερικό του σημείο. Άρα η $f : U \rightarrow V$ είναι ένα-προς-ένα στο U και επί του V , οπότε ορίζεται η αντίστροφη συνάρτηση $f^{-1} : V \rightarrow U$. Η f^{-1} είναι συνεχώς παραγωγίσιμη στο διάστημα V και ισχύει

$$(f^{-1})'(y) = \frac{1}{f'(x)} \quad \text{για κάθε } y \in V, \text{ όπου } x = f^{-1}(y) \in U.$$

Παρόμοιο αποτέλεσμα έχουμε όταν $f'(x_0) < 0$.

Θα δούμε τώρα πώς αυτό το αποτέλεσμα γενικεύεται για διανυσματικές συναρτήσεις με n συντεταγμένες πραγματικές συναρτήσεις, η καθεμία από τις οποίες έχει n μεταβλητές. Έχουμε το **Θεώρημα Αντίστροφης Απεικόνισης**.

Έστω διανυσματική συνάρτηση $\mathbf{f} : A \rightarrow \mathbb{R}^n$, όπου $A \subseteq \mathbb{R}^n$, και έστω $\mathbf{x}_0$ εσωτερικό σημείο του A και $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Υποθέτουμε ότι η $\mathbf{f}$ έχει συνεχείς μερικές παραγώγους πρώτης τάξης στο A και

$$J\mathbf{f}(\mathbf{x}_0) \neq 0.$$

Τότε υπάρχει περιοχή $U \subseteq A$ στον $\mathbb{R}^n$ με το $\mathbf{x}_0$ ως εσωτερικό της σημείο και περιοχή V στον $\mathbb{R}^n$ με το $\mathbf{y}_0$ ως εσωτερικό της σημείο έτσι ώστε η $\mathbf{f} : U \rightarrow V$ να είναι ένα-προς-ένα στην U και επί της V . Άρα ορίζεται η αντίστροφη συνάρτηση $\mathbf{f}^{-1} : V \rightarrow U$ η οποία έχει συνεχείς μερικές παραγώγους πρώτης τάξης στην V και ισχύει

$$D(\mathbf{f}^{-1})(\mathbf{y}) = (D\mathbf{f}(\mathbf{x}))^{-1}, \quad J(\mathbf{f}^{-1})(\mathbf{y}) = \frac{1}{J\mathbf{f}(\mathbf{x})} \quad \text{για κάθε } \mathbf{y} \in V, \text{ όπου } \mathbf{x} = \mathbf{f}^{-1}(\mathbf{y}) \in U.$$

Η απόδειξη βασίζεται στο Θεώρημα Πεπλεγμένης Συνάρτησης. Θεωρούμε τις συντεταγμένες συναρτήσεις $f_1, \dots, f_n : A \rightarrow \mathbb{R}$ της $\mathbf{f}$ και ορίζουμε τις συναρτήσεις $F_1, \dots, F_n : A \times \mathbb{R}^n \rightarrow \mathbb{R}$ με τους τύπους:

$$\begin{aligned} F_1(\mathbf{x}, \mathbf{y}) &= F_1(x_1, \dots, x_n, y_1, \dots, y_n) = f_1(x_1, \dots, x_n) - y_1 \\ &\vdots \\ F_n(\mathbf{x}, \mathbf{y}) &= F_n(x_1, \dots, x_n, y_1, \dots, y_n) = f_n(x_1, \dots, x_n) - y_n \end{aligned}$$

για κάθε $\mathbf{x} \in A, \mathbf{y} \in \mathbb{R}^n$.

Κατόπιν, θεωρούμε το σύνολο S των σημείων $(\mathbf{x}, \mathbf{y}) = (x_1, \dots, x_n, y_1, \dots, y_n) \in A \times \mathbb{R}^n$ τα οποία ικανοποιούν τις n εξισώσεις

$$\begin{aligned} F_1(\mathbf{x}, \mathbf{y}) &= f_1(x_1, \dots, x_n) - y_1 = 0 \\ &\vdots \\ F_n(\mathbf{x}, \mathbf{y}) &= f_n(x_1, \dots, x_n) - y_n = 0 \end{aligned}$$

Είναι σαφές ότι αυτές οι εξισώσεις ισοδυναμούν με την διανυσματική εξίσωση

$$\mathbf{f}(\mathbf{x}) = \mathbf{y}.$$

Δηλαδή το σύνολο S είναι το σύνολο των σημείων $(\mathbf{x}, \mathbf{y})$ με $\mathbf{x} \in A$ και $\mathbf{y} = \mathbf{f}(\mathbf{x}) \in \mathbb{R}^n$. Ισοδύναμα, το σύνολο S είναι το γράφημα της συνάρτησης $\mathbf{f}$:

$$S = \{(\mathbf{x}, \mathbf{f}(\mathbf{x})) \mid \mathbf{x} \in A\}.$$

Γνωρίζουμε ότι το σημείο $(\mathbf{x}_0, \mathbf{y}_0)$ ανήκει στο S , αφού $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$.

Επίσης, γνωρίζουμε ότι οι $F_1, \dots, F_n$ έχουν συνεχείς μερικές παραγώγους πρώτης τάξης ως προς όλες τις μεταβλητές σε μια περιοχή του σημείου $(\mathbf{x}_0, \mathbf{z}_0)$ και ότι στο σημείο αυτό ισχύει

$$\det \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(\mathbf{x}_0, \mathbf{y}_0) & \dots & \frac{\partial F_1}{\partial x_n}(\mathbf{x}_0, \mathbf{y}_0) \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(\mathbf{x}_0, \mathbf{y}_0) & \dots & \frac{\partial F_n}{\partial x_n}(\mathbf{x}_0, \mathbf{y}_0) \end{bmatrix} = \det \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}_0) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}_0) \end{bmatrix} = J\mathbf{f}(\mathbf{x}_0) \neq 0. \quad (74)$$

Από το Θεώρημα Πεπλεγμένης Συνάρτησης συνεπάγεται ότι υπάρχει περιοχή V στον $\mathbb{R}^n$ με το $\mathbf{y}_0$ ως εσωτερικό της σημείο και περιοχή U' στον $\mathbb{R}^n$ με το $\mathbf{x}_0$ ως εσωτερικό της σημείο ώστε για κάθε $\mathbf{y} \in V$ να υπάρχει ακριβώς ένα $\mathbf{x} \in U'$ ώστε το σημείο $(\mathbf{x}, \mathbf{y})$ να ανήκει στο S ή, ισοδύναμα, να

ικανοποιεί την $\mathbf{f}(\mathbf{x}) = \mathbf{y}$. Αν θεωρήσουμε το σύνολο $U = U' \cap \mathbf{f}^{-1}(V) = \{\mathbf{x} \in U' \mid \mathbf{f}(\mathbf{x}) \in V\}$, τότε μπορούμε να αποδείξουμε (όχι εδώ) ότι το U έχει το $\mathbf{x}_0$ ως εσωτερικό του σημείο. Είναι φανερό ότι η $\mathbf{f} : U \rightarrow V$ είναι ένα-προς-ένα στο U και επί του V και άρα ορίζεται η αντίστροφη συνάρτηση $\mathbf{f}^{-1} : V \rightarrow U$.

Η συνάρτηση $\mathbf{f}^{-1}$ έχει συνεχείς μερικές παραγώγους πρώτης τάξης και επειδή

$$(\mathbf{f}^{-1} \circ \mathbf{f})(\mathbf{x}) = \mathbf{x} \quad \text{για κάθε } \mathbf{x} \in U$$

από τον Κανόνα Αλυσίδας έχουμε ότι

$$D(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x})) D\mathbf{f}(\mathbf{x}) = D(\mathbf{f}^{-1} \circ \mathbf{f})(\mathbf{x}) = I_n \quad \text{για κάθε } \mathbf{x} \in U,$$

όπου I είναι ο μοναδιαίος $n \times n$ πίνακας, αυτός που έχει κάθε διαγώνιο στοιχείο του ίσο με 1 και κάθε άλλο στοιχείο του ίσο με 0. Άρα οι $n \times n$ πίνακες $D(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x}))$ και $D\mathbf{f}(\mathbf{x})$ είναι αντίστροφοι:

$$D(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x})) = (D\mathbf{f}(\mathbf{x}))^{-1} \quad \text{για κάθε } \mathbf{x} \in U.$$

Επειδή οι ορίζουσες αντίστροφων πινάκων είναι αντίστροφοι αριθμοί συνεπάγεται

$$J(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x})) = \frac{1}{J\mathbf{f}(\mathbf{x})} \quad \text{για κάθε } \mathbf{x} \in U.$$

Αν γράψουμε τις συντεταγμένες συναρτήσεις της $\mathbf{f} : U \rightarrow V$ ως

$$\begin{aligned} y_1 &= y_1(x_1, \dots, x_n) \\ &\vdots \\ y_n &= y_n(x_1, \dots, x_n) \end{aligned}$$

και τις συντεταγμένες συναρτήσεις της αντίστροφης συνάρτησης $\mathbf{f}^{-1} : V \rightarrow U$ ως

$$\begin{aligned} x_1 &= x_1(y_1, \dots, y_n) \\ &\vdots \\ x_n &= x_n(y_1, \dots, y_n) \end{aligned}$$

τότε η σχέση $D(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x})) = (D\mathbf{f}(\mathbf{x}))^{-1}$ για τις παραγώγους γράφεται

$$\begin{bmatrix} \frac{\partial x_1}{\partial y_1}(\mathbf{y}) & \dots & \frac{\partial x_1}{\partial y_n}(\mathbf{y}) \\ \vdots & & \vdots \\ \frac{\partial x_n}{\partial y_1}(\mathbf{y}) & \dots & \frac{\partial x_n}{\partial y_n}(\mathbf{y}) \end{bmatrix} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial y_1}{\partial x_n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial y_n}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial y_n}{\partial x_n}(\mathbf{x}) \end{bmatrix}^{-1}$$

όπου $\mathbf{y} = \mathbf{f}(\mathbf{x})$ ή, ισοδύναμα, $\mathbf{x} = \mathbf{f}^{-1}(\mathbf{y})$.

Ομοίως, η σχέση $J(\mathbf{f}^{-1})(\mathbf{f}(\mathbf{x})) = \frac{1}{J\mathbf{f}(\mathbf{x})}$ για τις Ιακωβιανές ορίζουσες γράφεται

$$\frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)}(\mathbf{y}) = 1 / \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}(\mathbf{x}).$$

Παράδειγμα: Θεωρούμε την συνάρτηση από τον $\mathbb{R}^2$ στον $\mathbb{R}^2$ με τύπο

$$\begin{aligned} z &= z(x, y) = x^2 - y^2 \\ w &= w(x, y) = 3xy^3 \end{aligned} \tag{75}$$

Με αυτό εννοούμε ότι η συνάρτηση έχει συντεταγμένες συναρτήσεις τις z, w η καθεμία από τις οποίες είναι συνάρτηση των x, y .

Μας ενδιαφέρει να δούμε αν η συνάρτηση είναι αντιστρέψιμη κοντά στο σημείο $(x_0, y_0) = (-1, 1)$ και στο αντίστοιχο σημείο $(z_0, w_0) = (0, -3)$.

Κατ' αρχάς παρατηρούμε ότι οι συναρτήσεις z, w έχουν συνεχείς μερικές παραγώγους πρώτης τάξης ως προς x, y σε κάθε σημείο του $\mathbb{R}^2$. Η Ιακωβιανή ορίζουσα είναι ίση με

$$\frac{\partial(z, w)}{\partial(x, y)} = \det \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix} = \det \begin{bmatrix} 2x & -2y \\ 3y^3 & 9xy^2 \end{bmatrix} = 18x^2y^2 + 6y^4$$

και, ειδικότερα στο σημείο $(x_0, y_0) = (-1, 1)$ είναι

$$\frac{\partial(z, w)}{\partial(x, y)}(-1, 1) = 24 \neq 0.$$

Άρα η συνάρτηση είναι τοπικά αντιστρέψιμη. Συγκεκριμένα, υπάρχει περιοχή U με το $(x_0, y_0) = (-1, 1)$ ως εσωτερικό της σημείο και περιοχή V με το $(z_0, w_0) = (0, -3)$ ως εσωτερικό της σημείο έτσι ώστε για κάθε $(x, y) \in U$ υπάρχει μοναδικό $(z, w) \in V$ τα οποία συνδέονται με τις σχέσεις (75) και, αντιστρόφως, για κάθε $(z, w) \in V$ υπάρχει μοναδικό $(x, y) \in U$ τα οποία συνδέονται με τις αντίστροφες σχέσεις

$$\begin{aligned} x &= x(z, w) \\ y &= y(z, w) \end{aligned}$$

Τους ακριβείς τύπους των αντίστροφων σχέσεων ίσως να μην μπορούμε να τους βρούμε. Αλλά γνωρίζουμε ότι οι συναρτήσεις $x(z, w), y(z, w)$ έχουν συνεχείς μερικές παραγώγους πρώτης τάξης στην περιοχή V και ότι

$$\frac{\partial(x, y)}{\partial(z, w)}(z, w) = 1 / \frac{\partial(z, w)}{\partial(x, y)}(x, y) = 1 / (18x^2y^2 + 6y^4).$$

Π.χ. στο σημείο $(z_0, w_0) = (0, -3)$ είναι

$$\frac{\partial(x, y)}{\partial(z, w)}(0, -3) = 1 / \frac{\partial(z, w)}{\partial(x, y)}(-1, 1) = 1/24.$$

Επίσης,

$$\begin{aligned} \begin{bmatrix} \frac{\partial x}{\partial z}(z, w) & \frac{\partial x}{\partial w}(z, w) \\ \frac{\partial y}{\partial z}(z, w) & \frac{\partial y}{\partial w}(z, w) \end{bmatrix} &= \begin{bmatrix} \frac{\partial z}{\partial x}(x, y) & \frac{\partial z}{\partial y}(x, y) \\ \frac{\partial w}{\partial x}(x, y) & \frac{\partial w}{\partial y}(x, y) \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 2x & -2y \\ 3y^3 & 9xy^2 \end{bmatrix}^{-1} = \frac{1}{18x^2y^2 + 6y^4} \begin{bmatrix} 9xy^2 & 2y \\ -3y^3 & 2x \end{bmatrix} \end{aligned}$$

και άρα

$$\begin{aligned} \frac{\partial x}{\partial z}(z, w) &= \frac{9xy^2}{18x^2y^2 + 6y^4}, & \frac{\partial x}{\partial w}(z, w) &= \frac{2y}{18x^2y^2 + 6y^4}, \\ \frac{\partial y}{\partial z}(z, w) &= -\frac{3y^3}{18x^2y^2 + 6y^4}, & \frac{\partial y}{\partial w}(z, w) &= \frac{2x}{18x^2y^2 + 6y^4}. \end{aligned}$$