Δ. Χρήστου-Βαρσακέλης
Τμήμα Εφαρμοσμένης Πληροφορικής,
Πανεπιστήμιο Μακεδονίας
dcv@uom.gr

# Επιχειρησιακή Έρευνα

## Σημειώσεις Μαθήματος (Μη-γραμμικός Προγραμματισμός)

# Περιεχόμενα

# 1

# Εισαγωγική Σημείωση

Η Επιχειρησιακή Έρευνα (ΕΕ) αφορά στη χρήση και ανάπτυξη μαθηματικών και υπολογιστικών μεθόδων με σκοπό την επίλυση προβλημάτων βελτιστοποίησης (συνήθως), με εφαρμογές στην Πληροφορική, τη Διοίκηση Επιχειρήσεων, τις Επιστήμες Μηχανικού, τις Οικονομικές και Χρηματοοικονομικές επιστήμες, κ.α. Στις σημειώσεις αυτές παρουσιάζονται μερικές από τις κύριες αναλυτικές μεθόδους που εφαρμόζονται στην ΕΕ, και αφορούν τη μή γραμμική βελτιστοποίηση.

# Γραμμικοί Διανυσματικοί Χώροι

Τα παρακάτω είναι μια σύντομη αναδρομή σε μερικές βασικές έννοιες από την μαθηματική ανάλυση. Για λεπτομέρειες βλ. Rudin, "Principles of Mathematical Analysis" και Strang, "Linear Algebra and its Applications".

**Ορισμός 0.1**

να σύνολο $V$ είναι πραγματικός διανυσματικός χώρος (ΠΔΧ) αν $\forall$ (για όλα τα) $x, y, z \in V$ και $a, b \in \mathbb{R}$, ισχύουν τα παρακάτω:

1. $x + y \in V$
2. $ax \in V$ (κλειστό στην πρόσθεση και πολλαπλασιασμό με αριθμ. μεταβλητή)

και

1. $(x + y) + z = x + (y + z)$
2. $0 + x = x$
3. $x + (-x) = 0$
4. $x + y = y + x$
5. $a(x + y) = ax + ay$
6. $(a + b)x = ax + bx$
7. $(ab)x = a(bx)$
8. $1 \cdot x = x$

Παραδείγματα ΠΔΧ συμπεριλαμβάνουν το σύνολο των πραγματικών, το σύνολο $n$-διάστατων πραγματικών διανυσμάτων $\mathbb{R}^n$, πραγματικούς $m \times n$ πίνακες, $\mathbb{R}^{m\times n}$, και το σύνολο των συνεχών, πραγματικών συναρτήσεων ορισμένων σε ένα υποσύνολο $[t_0, t_1]$ του $\mathbb{R}$ (γράφεται $C[t_0, t_1]$). Βεβαιωθείτε ότι μπορείτε να δείξετε ότι κάθε ένα από τα παραπάνω είναι πράγματι διανυσματικός χώρος.

## 2.1 Βάση, ανεξαρτησία και διάσταση

Ένα σύνολο διανυσμάτων $v_1, ..., v_n \in V$ είναι (γραμμικά) **ανεξάρτητα** αν

$$c_1 v_1 + c_2 v_2 + \ldots c_n v_n = 0, \qquad c_1, ..., c_n \in \mathbb{R}$$

μόνο όταν $c_1 = c_2 = ... = c_n = 0$. Αν τα $v_1, ..., v_n$ δεν είναι γραμμικά ανεξάρτητα, τότε είναι γραμμικά εξαρτημένα και ένα (ή περισσότερα) από αυτά είναι γραμμικός των υπολοίπων.

**Βάση** ενός διανυσματικού χώρου $V$ είναι ένα σύνολο διανυσμάτων (στοιχεία του χώρου $V$) τα οποία έχουν τις εξής ιδιότητες: i) είναι γραμμικά ανεξάρτητα, και ii) κάθε στοιχείο του $V$ μπορεί να εκφραστεί σαν γραμμικός συνδυασμός των διανυσμάτων της βάσης. Οι βάσεις *δεν είναι* μοναδικές. Παράδειγμα: Συνήθως στο $\mathbb{R}^3$ χρησιμοποιείται η ορθοκανονική βάση $e_1 = [1\ 0\ 0]^T$, $e_2 = [0\ 1\ 0]^T$ και $e_3 = [0\ 0\ 1]^T$. Η **διάσταση** ενός γραμμικού διανυσματικού χώρου $V$ είναι ο αριθμός των (ανεξάρτητων) διανυσμάτων τα οποία σχηματίζουν βάση για τον $V$.

Σημειώστε ότι ένας διανυσματικός χώρος *δεν* χρειάζεται να έχει απαραίτητα πεπερασμένο αριθμό διαστάσεων. Για παράδειγμα, ο χώρος των συναρτήσεων $C[t_0, t_1]$ (βλ. προηγούμενη παράγραφο) είναι απειροδιάστατος: Γενικά, μία συνεχής συνάρτηση $f(t)$ για $t_0 \leq t \leq t_1$ *δεν* μπορεί να εκφραστεί σαν γραμμικός συνδυασμός οποιουδήποτε πεπερασμένου αριθμού συνεχών συναρτήσεων ορισμένων στο ίδιο χρονικό διάστημα.

## 2.2 Γραμμικοί τελεστές - γραμμική απεικόνιση (Linear operators)

Εν συντομία, *γραμμικοί τελεστές* είναι μαθηματικά "αντικείμενα" τα οποία "δρουν" πάνω σε διανύσματα ενός δεδομένου διανυσματικού χώρου (πεδίο ορισμού) και "παράγουν" άλλα διανύσματα, πιθανόν σε διαφορετικό διανυσματικό χώρο (πεδίο τιμών). Ακριβέστερα:

Ο $L$ είναι γραμμικός τελεστής από τον (γρ. διαν.) χώρο $X$ στον χώρο $Y$ αν για κάθε $x_1, x_2 \in X$ και για κάθε $a \in \mathbb{R}$, ισχύουν:

$$\begin{aligned} &L(x_1) \in Y \\ &L(x_1 + x_2) = L(x_1) + L(x_2) \\ &L(ax_1) = aL(x_1) \end{aligned}$$

Συνήθως γράφουμε $L : X \to Y$ και λέμε ότι ο "$L$ στέλνει ή απεικονίζει το $X$ στο $Y$". Γνωρίζετε ήδη μια συνηθισμένη τάξη γραμμικών τελεστών, τους πίνακες. Για παράδειγμα, ένας πραγματικός πίνακας διαστάσεων $3 \times 5$ είναι γραμμικός τελεστής που στέλνει 5-διάστατα διανύσματα σε 3-διάστατα διανύσματα, μέσω πολλαπλασιασμού από τα αριστερά. Αν $x \in \mathbb{R}^5$, $y \in \mathbb{R}^3$, τότε μπορούμε να γράψουμε: $M : \mathbb{R}^5 \to \mathbb{R}^3$ και $M \cdot x = y$, όπου "$\cdot$" δηλώνει το συνηθισμένο πολλαπλασιασμό πινάκων.

Ο αναγνώστης θα πρέπει να βεβαιωθεί ότι έχει σχετική άνεση με τα βασικά στοιχεία γραμμικής άλγεβρας, συμπεριλαμβανομένων του πολλαπλασιασμού πινάκων, και τις κύριες ιδιότητες πινάκων όπως **βαθμός, χώρος στηλών (range space), και μηδενοχώρος (nullspace), ορίζουσα (determinant), ιδιοδιανύσματα και ιδιοτιμές (eigenvectors/eigenvalues)**. Βλέπε επίσης Παράρτημα Α΄.

## 2.3 Εσωτερικό Γινόμενο

Το εσωτερικό γινόμενο δύο διανυσμάτων (π.χ. $v_1^T v_2 \in \mathbb{R}$ για $v_1, v_2 \in \mathbb{R}^3$) μπορεί άμεσα να γενικευτεί ως εξής: Με δεδομένα δύο στοιχεία ενός διανυσματικού χώρου, το εσωτερικό γινόμενο είναι ένας τρόπος να συνδιαστούν αυτά ώστε να προκύψει ένας πραγματικός αριθμός. Πιό συγκεκριμένα, έστω ότι $X$ είναι ένας γραμμικός διανυσματικός χώρος, και $x_1, x_2 \in X$, $a, b \in \mathbb{R}$. Λέμε ότι ο τελεστής $\langle x_1, x_2 \rangle$ είναι ένα θετικά ορισμένο) **εσωτερικό γινόμενο** αν:

- $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{R}$
- $\langle x_1, x_2 \rangle = \langle x_2, x_1 \rangle$ (συμμετρική ιδιότητα)
- $\langle x_1, x_1 \rangle \geq 0$ (μη-αρνητικότητα)
- $\langle x_1, x_1 \rangle = 0$ αν και μόνο αν $x_1 = 0$
- $\langle x_1, ax_2 + bx_3 \rangle = a\langle x_1, x_2 \rangle + b\langle x_1, x_3 \rangle$

Για να θυμόμαστε σε ποιό διανυσματικό χώρο αναφέρεται ένα εσωτερικό γινόμενο, θα γράφουμε $\langle \cdot, \cdot \rangle_X$ για το εσωτερικό γινόμενο το οποίο ορίζεται στο χώρο $X \times X$. Τα παρακάτω είναι παραδείγματα εσωτερικών γινομένων:

- Το σύνηθες εσωτερικό γινόμενο στο $X = \mathbb{R}^3$: Για ένα ζεύγος διανυσμάτων $x, y$ στο $\mathbb{R}^3$, έχουμε $\langle x, y \rangle_{\mathbb{R}^3} = x^T y = x_1 y_1 + x_2 y_2 + x_3 y_3$.
- Έστω $X = \mathbb{R}^{m \times n}$, ο χώρος των πραγματικών $m \times n$ πινάκων (όπου $m, n$ θετικοί ακέραιοι). Για δύο πίνακες $A, B \in X$ μπορούμε να ορίσουμε:

  $$\langle A, B \rangle_{\mathbb{R}^{m \times n}} = tr(A^T B) = \sum_{i=1}^{n} \sum_{j=1}^{m} A_{ij} B_{ij}$$

  όπου $tr$ συμβολίζει το **ίχνος** ενός *τετράγωνου* πίνακα, το οποίο ορίζεται ως το άθροισμα των διαγώνιων στοιχείων του πίνακα. Αν ο $A$ είναι πίνακας $n \times n$, τότε $tr(A) \triangleq \sum_{i=1}^{n} A_{ii}$.
- Παρατηρούμε ότι ο χώρος $X$ δεν χρειάζεται να έχει πεπερασμένη διάσταση. Αν $f, g$ είναι δύο πραγματικές, συνεχείς συναρτήσεις πραγματικής μεταβλητής ορισμένες στο διάστημα $[0, T]$, (δηλ. $f, g \in X = C[0, T]$), τότε μπορούμε να ορίσουμε

  $$\langle f, g \rangle_{C[0,T]} = \int_0^T f(t) g(t) dt$$

# 3

## Μη-Γραμμικός Προγραμματισμός

Ξεκινάμε με την ανασκόπηση ορισμένων βασικών προτάσεων από την ανάλυση. Η κατανόηση των παρακάτω, προϋποθέτει άνεση με το υλικό των παραρτημάτων Α' (Στοιχεία Γραμμικής Άλγεβρας) και Β' (Παραγώγιση συναρτήσεων διανυσματικής μεταβλητής) στο τέλος αυτών των σημειώσεων.

### 3.1 Βελτιστοποίηση βαθμωτής συνάρτησης πραγματικής μεταβλητής

**Θεώρημα 1.1**

Έστω η συνάρτηση $\psi : \mathbb{R} \to \mathbb{R}$ συνεχής και παραγωγίσιμη. Άν το $min_x \; \psi(x)$ υπάρχει, έστω στο σημείο $x_0$, τότε

$$\left.\frac{d\psi}{dx}\right|_{x_0} = 0, \qquad \left.\frac{d^2\psi}{dx^2}\right|_{x_0} > 0$$

Απόδειξη: Είναι γνωστό από την μαθηματική ανάλυση (Θεώρημα Taylor) ότι αν η συνάρτηση $f : \mathbb{R} \to \mathbb{R}$ είναι απείρως παραγωγίσιμη, τότε η $f(x)$ μπορεί να εκφραστεί σαν γραμμικός συνδιασμός των παραγώγων της, υπολογισμένων σε οποιοδήποτε σημείο $x_0$ που ανήκει στο πεδίο ορισμού της $f$:

$$f(x) = f(x_0) + \left.\frac{df}{dx}\right|_{x_0}(x - x_0) + \frac{1}{2}\left.\frac{d^2 f}{dx^2}\right|_{x_0}(x - x_0)^2 + ... = \sum_{0}^{\infty}\frac{1}{i!}\left.\frac{d^i f}{dx^i}\right|_{x_0}(x - x_0)^i$$

Η δεξιά πλευρά της παραπάνω σχέσης είναι γνωστή και ως *σειρά Taylor*.

Γράφουμε λοιπόν την σειρά Taylor για την $\psi$ στο σημείο $x_0$:

$$\psi(x_0 + \delta) = \psi(x_0) + \left.\frac{d\psi}{dx}\right|_{x_0}\delta + \frac{1}{2}\frac{d^2\psi}{dx^2}\delta^2 + \text{ο.μ.τ.}$$

όπου "ο.μ.τ" δηλώνει όρους μεγαλύτερων τάξεων (τρίτης και άνω) ως προς . Υποθέτουμε οτι ενώ το $x_0$ είναι τοπικό ελάχιστο, χουμε $d\psi/dx \neq 0$. Σε αυτή την περίπτωση, επιλέγουμε $\delta = -\epsilon(d\psi/dx)$ όπου $\epsilon > 0$, το οποίο μας δίνει:

$$\psi(x_0+\delta) = \psi(x_0) - \epsilon\left(\left.\frac{d\psi}{dx}\right|_{x_0}\right)^2 + \epsilon^2\left(\frac{1}{2}\left(\left.\frac{d^2\psi}{dx^2}\right|_{x_0}\right)\left(\left.\frac{d\psi}{dx}\right|_{x_0}\right)^2\right) + \text{ο.μ.τ.}$$

Έστω τώρα ότι το $\epsilon$ γίνεται ολοένα και μικρότερο. Επειδή $\epsilon^2 \to 0$ "γρηγορότερα" από ότι $\epsilon \to 0$, θα υπάρχει τιμή του $\epsilon > 0$ τέτοια ώστε ο γραμμικός όρος να κυριαρχεί (δηλ. να είναι μεγαλύτερος - κατ'απόλυτη τιμή - από όλους τους επόμενους όρους *μαζί*), άρα $\psi(x_0+\delta) < \psi(x_0)$. Αλλά τότε το $x_0$ δεν μπορεί είναι τοπικό ελάχιστο όπως υποθέσαμε αρχικά (άτοπο).

Στη συνέχεια, θέτοντας $\frac{dx}{dt}|_{x_0} = 0$, βλέπουμε ότι το πρόσημο του $\psi(x_0+\delta) - \psi(x_0)$ καθορίζεται από τον όρο $\frac{d^2\psi}{dx^2}$ όταν το $\epsilon$ είναι επαρκώς μικρό, και άρα η δεύτερη παράγωγος του $\psi$ πρέπει να είναι μη αρνητική στο τοπικό ελάχιστο. □

## 3.2 Βελτιστοποίηση συναρτήσεων διανυσματικής μεταβλητής

Για πραγματικές (βαθμωτές) συναρτήσεις διανυσματικής μεταβλητής, η κατάσταση είναι βασικά "παράλληλη" με αυτήν που αναφέραμε παραπάνω. Η μόνη διαφορά είναι ότι για την $\psi : \mathbb{R}^n \to \mathbb{R}$, η ποσότητα $\frac{d\psi}{dx}$ πρέπει να αντικατασταθεί από ένα διάνυσμα (μερική παραγώγιση ως προς την κάθε μεταβλητή της $\psi$), ενώ η $\frac{d^2\psi}{dx^2}$ πρέπει να αντικατασταθεί από όλες τις δεύτερες μερικές παραγώγους της $\psi$ (τις οποίες θα είναι βολικό να οργανώσουμε σε έναν τετραγωνικό πίνακα). Έτσι, θα πρέπει να αντικαταστήσουμε την ποσότητα $\frac{d\psi}{dx}\delta$ με την $\langle\frac{d\psi}{dx}, \delta\rangle \triangleq \left(\frac{d\psi}{dx}\right)^T \delta$ και την $\frac{d^2\psi}{dx^2}\delta^2$ με $\delta^T\left(\frac{d^2\psi}{dx^2}\right)\delta$ στην απόδειξη του Θεωρήματος 1.1.

**Θεώρημα 2.1**

Έστω η συνάρτηση $\psi : \mathbb{R}^n \to \mathbb{R}$ συνεχής και παραγωγίσιμη. Εάν το $min\ \psi = x_0$ υπάρχει, τότε

$$\left.\frac{\partial\psi}{\partial x}\right|_{x_0} \triangleq \begin{bmatrix} \partial\psi/\partial x_1 \\ \partial\psi/\partial x_2 \\ \vdots \\ \partial\psi/\partial x_n \end{bmatrix} = 0$$

και

$$\frac{\partial^2\psi}{\partial x^2} \triangleq \begin{bmatrix} \frac{\partial\psi}{\partial x_1\partial x_1} & \frac{\partial\psi}{\partial x_1\partial x_2} & \cdots & \frac{\partial\psi}{\partial x_1\partial x_n} \\ \frac{\partial\psi}{\partial x_2\partial x_1} & \frac{\partial\psi}{\partial x_2\partial x_2} & \cdots & \\ \vdots & & & \\ \frac{\partial\psi}{\partial x_n\partial x_1} & \frac{\partial\psi}{\partial x_n\partial x_2} & \cdots & \frac{\partial\psi}{\partial x_n\partial x_n} \end{bmatrix} > 0$$

Απόδειξη: Έστω $\delta \in \mathbb{R}^n$ διάνυσμα του οποίου την τιμή μπορούμε να επιλέξουμε, τότε

$$\psi(x_0+\delta) = \psi(x_0) + \langle \left.\frac{\partial \psi}{\partial x}\right|_{x_0}, \delta\rangle + \frac{1}{2}\delta^T \frac{\partial^2 \psi}{\partial x^2}\delta + \text{ο.μ.τ.}$$

Έστω ότι $x_0$ είναι τοπικό ελάχιστο της $\psi$, αλλά $d\psi/dx \neq 0$, τότε θέτουμε $\delta = -\epsilon(d\psi/dx)$, το οποίο μας δίνει:

$$\psi(x_0+\delta) = \psi(x_0) - \epsilon \left( \left.\frac{\partial \psi}{\partial x}\right|_{x_0} \right)^T \left.\frac{\partial \psi}{\partial x}\right|_{x_0} + \epsilon^2 \left( \frac{1}{2} \left( \left.\frac{\partial \psi}{\partial x}\right|_{x_0} \right)^T \frac{\partial^2 \psi}{\partial x^2} \left( \left.\frac{\partial \psi}{\partial x}\right|_{x_0} \right) \right) + \ldots$$

Όπως και στο Θεώρημα 1.1, παρατηρούμε ότι για τιμές του $\epsilon$ επαρκώς μικρές, η δεξιά πλευρά της εξίσωσης μπορεί να καταστεί αρνητική, άρα το $x_0$ δεν μπορεί να είναι τοπικό ελάχιστο (άτοπο). $\square$

Παρατηρούμε ότι ο πίνακας $\frac{\partial^2 \psi}{\partial x \partial x}$ είναι συμμετρικός διότι $\frac{\partial \psi}{\partial x_i \partial x_j} = \frac{\partial \psi}{\partial x_j \partial x_i}$ (η σειρά με την οποία λαμβάνονται οι μερικές παράγωγοι δεν αλλάζει το αποτέλεσμα). Επίσης, η ικανή συνθήκη για τ.ελάχιστο στο $x_0$ είναι

$$\delta^T \frac{\partial^2 \psi}{\partial x \partial x}\delta > 0 \quad \forall\, \delta \in \mathbb{R}^n$$

Όταν ισχύει η παραπάνω ανισότητα για όλα τα (μη-μηδενικά) διανύσματα $\delta$, τότε γράφουμε

$$\frac{\partial^2 \psi}{\partial x \partial x} > 0$$

και λέμε ότι ο πίνακας $H = \frac{\partial^2 \psi}{\partial x \partial x}$ των δεύτερων μερικών παραγώγων της $\psi$ (γνωστός και σαν Εσσιανός πίνακας) είναι *θετικά ορισμένος* (βλ. Παράρτημα Α.5').

### 3.2.1 Σύνοψη

- Για να ελαχιστοποιήσουμε (ή να μεγιστοποιήσουμε) μια συνάρτηση $\psi(x)$ (όπου το $x$ ανήκει σε πεπερασμένων διαστάσεων διανυσματικό χώρο), εκφράζουμε την $\psi(x)$ σαν σειρά Taylor γύρω από ένα σημείο το οποίο υποτίθεται ότι είναι βέλτιστο (τ. ελάχιστο ή τ. μέγιστο).
- Για να έχει η συνάρτηση τοπικό βέλτιστο θα πρέπει **όλες** οι **περιττές παράγωγοι (1η, 3η, 5η,...)** που εμφανίζονται στη σειρά Taylor να μηδενίζονται, **μέχρι να συναντήσουμε μια μη-μηδενική άρτια** παράγωγο. Το πρόσημο αυτής της άρτιας παραγώγου καθορίζει αν το τ. βέλτιστο είναι μέγιστο ή ελάχιστο.
  Ετσι, για παράδειγμα η $\psi(x) = x^3$ **δεν έχει** βέλτιστο στο $\mathbb{R}$ (η πρώτη παράγωγος μηδενίζεται μόνο στο 0 αλλά σε εκείνο το σημείο δεν υπάρχει άρτια μη-μηδενική παράγωγος), ενώ η $\psi(x) = x^4$ **έχει** ελάχιστο στο 0 (η πρώτη παράγωγος μηδενίζεται στο 0 αλλά η δεύτερη μηδενίζεται και αυτή, συνεχίζοντας όμως βλέπουμε ότι και η 3η παράγωγος μηδενίζεται στο 0 ενώ η 4η είναι θετική στο ίδιο σημείο - οπότε συμπεραίνουμε ότι έχουμε ελάχιστο στο σημείο 0).

- Στην περίπτωση πολυμεταβλητής συνάρτησης, αν ο πίνακας των δεύτερων παραγώγων $H = \frac{\partial^2 \psi}{\partial x \partial x}$ είναι θετικά ορισμένος (στο σημείο που μηδενίζεται η $\partial\psi/\partial x$) τότε έχουμε τοπικό ελάχιστο, ενώ αν $H < 0$ (αρνητικά ορισμένος) έχουμε τοπικό μέγιστο. Αν ο $H$ δεν είναι ούτε θετικά ούτε αρνητικά ορισμένος (π.χ. έχει και θετικές και αρνητικές ιδιοτιμές - βλ. Παράρτημα Α5) τότε προκειται για σαγματικό σημείο (saddle point).
- Τέλος, αν ο $H$ είναι, για παράδειγμα, θετικά *ημι*-ορισμένος, και όχι αυστηρά θετικά ορισμένος όπως απαιτεί η ικανή συνθήκη για τ. ελάχιστο, θα μπορούμε να συμπεράνουμε ότι πράγματι βρισκόμαστε σε ελάχιστο της $\psi$, αν η $\psi$ είναι δευτέρου βαθμού (συνεπώς η σειρά Taylor της $\psi$ "σταματά" στον όρο που αφορά την δεύτερη παράγωγο, και δεν υπάρχουν όροι μεγαλύτερης τάξης οι οποίοι θα μπορούσαν να αλλάξουν το πρόσημο της διαφοράς $\psi(x_0 + \delta) - \psi(x_0)$. Σε περίπτωση που η $\psi$ είναι τρίτου ή μεγαλύτερου βαθμού, η συνθήκη $H \geq 0$ δε μας επιτρέπει να συμπεράνουμε ότι έχουμε τ. ελάχιστο χωρίς να εξετάσουμε τα πρόσημα των όρων μεγαλύτερων τάξεων στη σειρά Taylor της $\psi$.

## 3.3 Βελτιστοποίηση με Περιορισμούς

Η μέθοδος των **πολλαπλασιαστών Lagrange** επιτρέπει να ελαχιστοποιήσουμε (μεγιστοποιήσουμε) συναρτήσεις υπό περιορισμούς στις μεταβλητές των. Θα εξετάσουμε πρώτα την περίπτωση που οι περιορισμοί έχουν τη μορφή ισοτήτων.

### 3.3.1 Βελτιστοποίηση με περιορισμούς ισότητας

Έστω οι συναρτήσεις $\psi, \phi_1, ..., \phi_m : \mathbb{R}^n \to \mathbb{R}$ παραγωγίσιμες. Όπως έχουμε δει και στο μάθημα της Μαθηματικής Ανάλυσης, για να ελαχιστοποιήσουμε την $\psi(x)$ υπό τους περιορισμούς (υ.π.) $\phi_1(x) = \phi_2(x) = ... = \phi_m(x) = 0$, σχηματίζουμε την Λαγκραντζιανή (Lagrangian)

$$L = \psi(x) + \lambda_1 \phi_1(x) + \lambda_2 \phi_2(x) + ... + \lambda_m \phi_m(x)$$

ή, χρησιμοποιώντας τα διανύσματα $\lambda = [\lambda_1, ..., \lambda_m]^T, \phi = [\phi_1, ..., \phi_m]^T$:

$$L = \psi(x) + \lambda^T \phi \tag{3.1}$$

Στη συνέχεια, η συνθήκη

$$\left.\frac{\partial L}{\partial x}\right|_{x_0} = 0 \tag{3.2}$$

χρησιμοποιήθηκε για να εντοπίσουμε τα πιθανά τοπικά ελάχιστα, $x_0$, της συνάρτησης $\psi$ τα οποία ικανοποιούν και τις απαιτούμενες συνθήκες $\phi_i(x_0) = 0$.

Σε αντίθεση όμως με την περίπτωση όπου η $\psi$ βελτιστοποιείται χωρίς περιορισμούς, η συνθήκη $\left.\frac{\partial L}{\partial x}\right|_{x_0} = 0$ δεν αρκεί από μόνη της για να προσδιορίσουμε το $x_0$. Για παράδειγμα, εάν

$$\psi(x) = x_1 + x_2, \qquad \phi_1 = x_1^2 + x_2^2 = 0$$

τότε

$$L = x_1 + x_2 + \lambda(x_1^2 + x_2^2).$$

και συνθήκη πρώτης τάξης είναι

$$\frac{\partial L}{\partial x} = [1 + 2\lambda x_1, 1 + 2\lambda x_2]^T = 0.$$

Παρατηρούμε ότι η $\psi$ ελαχιστοποιείται στο $x_1 = x_2 = 0$ (γιατί αυτό είναι το μόνο εφικτό σημείο, λόγω της ιδιαίτερης μορφής του περιορισμού $\phi_1$), αλλά στο σημείο αυτό

$$\frac{\partial L}{\partial x} = [1\ 1]^T \neq 0 \text{ !!!}$$

Φαίνεται επομένως ότι στο συγκεκριμένο πρόβλημα η συνθήκη $\partial L/\partial x = 0$ δεν "εντοπίζει" το βέλτιστο.

Για να αποφεύγονται "παθολογικές" περιπτώσεις, όπως η παραπάνω, χρειαζόμαστε κάποιες πρόσθετες προϋποθέσεις τις οποίες πρέπει να πληρούν οι περιορισμοί $\phi_i$. Το σημείο στο οποίο πρέπει να εστιάσουμε είναι η γεωμετρική ερμηνεία της συνθήκης

$$\frac{\partial L}{\partial x} = 0 \Rightarrow \frac{\partial \psi}{\partial x} + \sum_1^m \lambda_i \phi_x = 0 \Rightarrow \frac{\partial \psi}{\partial x} = -\sum_1^m \lambda_i \phi_x.$$

Η τελευταία πρόταση απαιτεί **να μπορεί η κλίση της αντικειμενικής συνάρτησης, $\psi$, να εκφραστεί ως γραμμικός συνδυασμός των κλίσεων των περιορισμών, στο βέλτιστο σημείο**.

Επανεξετάζοντας το παράδειγμά στο οποίο η συνθήκη $\partial L/\partial x = 0$ "αποτυγχάνει":

$$\psi = x_1 + x_2, \qquad \phi_1 = x_1^2 + x_2^2 = 0$$

βλέπουμε ότι η κλίση (gradient) της $\psi$ είναι

$$\frac{d\psi}{dx} = [1,\ 1]^T$$

ενώ η παράγωγος του περιορισμού είναι

$$\frac{d\phi}{dx} = [2x_1,\ 2x_2]^T$$

Άρα η παραβίαση της συνθήκης $\partial L/\partial x = 0$ στο ελάχιστο $(0,0)$ συμπίπτει με το γεγονός ότι η $\psi$ ελαχιστοποιείται στο σημείο όπου η παράγωγος του περιορισμού μηδενίζεται, και άρα είναι αδύνατο να εκφράσουμε την $\partial f/\partial x$ ως πολλαπλάσιο της κλίσης του μοναδικού περιορισμού. Τα παραπάνω συνοψίζονται στο εξής θεώρημα:

**Θεώρημα 3.1**

Έστω $\psi : \mathbb{R}^n \to \mathbb{R}$, και $\phi_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, ..., m$ ώστε ο πίνακας

$$\left[\frac{\partial \phi_1}{\partial x}, \frac{\partial \phi_2}{\partial x}, ..., \frac{\partial \phi_m}{\partial x}\right]$$

να έχει σταθερή τάξη $k$ σε όλα τα σημεία που ανήκουν σε γειτονιά του συνόλου

$$S : \{x : \phi_1(x) = ... = \phi_m(x) = 0\}$$

Τότε, για $L = f + \lambda^T \phi(x)$, όπου $\lambda = [\lambda_1, ..., \lambda_m]^T$ και $\phi(x) = [\phi_1(x), ..., \phi_m(x)]^T$, η συνθήκη

$$\left.\frac{\partial L}{\partial x}\right|_{x_0} = 0 \tag{3.3}$$

είναι αναγκαία ωστε το $x_0$ να είναι τοπικό βέλτιστο της $\psi$ υ.π. $\phi_i = 0$. Στο $x_0$ έχουμε:

$$\left.\frac{\partial \psi}{\partial x}\right|_{x_0} = \sum_{i=1}^{m} \lambda_i \frac{\partial \phi_i}{\partial x}$$

Απόδειξη: Αναπτύσουμε τις $\psi$ και $\phi_i$ χρησιμοποιόντας τις σειρές Taylor:

$$\psi(x_0 + \delta) = \psi(x_0) + \frac{d\psi}{dx}(\delta) + \frac{1}{2}\frac{d^2\psi}{dx^2}(\delta, \delta) + ... \tag{3.4}$$

$$\phi_i(x_0 + \delta) = \phi_i(x_0) + \frac{d\phi_i}{dx}(\delta) + \frac{1}{2}\frac{d^2\phi_i}{dx^2}(\delta, \delta) + ... \tag{3.5}$$

όπου για συντομία γράφουμε $\frac{d\psi}{dx}(\delta)$ αντί για $\left(\frac{d\psi}{dx}\right)^T \delta$, και $\frac{d^2\psi}{dx^2}(\delta, \delta)$ αντί για $\delta^T \frac{d^2\psi}{dx^2} \delta$.

Παρατηρούμε ότι λόγω ύπαρξης περιορισμών $\phi(x_0) = 0$, δεν είναι εφικτές όλες οι "διαταράξεις" $\delta$: συγκεκριμένα, κάθε $\delta$ που προστίθεται στο $x_0$ πρέπει να το "μετακινεί" έτσι ώστε να συνεχίζουν να ισχύουν οι περιορισμοί, δηλαδή $\phi_i(x_0 + \delta) = 0$. Αυτό θα συμβαίνει όταν (για αρκετά μικρό $\|\delta\|$) $\left(\frac{\partial \phi_i}{\partial x}\right)^T \delta = 0$, για κάθε $i$. Βλέπουμε λοιπόν ότι οι αποδεκτές επιλογές για το $\delta$ είναι μόνο αυτές για τις οποίες

$$\left[\frac{\partial \phi_1}{\partial x}, \frac{\partial \phi_2}{\partial x}, ..., \frac{\partial \phi_m}{\partial x}\right]^T \cdot \delta = 0$$

ή, ισοδύναμα, το  πρέπει να ανήκει στο μηδενοχώρο του πίνακα που ορίζεται από τις παραγώγους (κλίσεις) των περιορισμών:

$$\delta \in Ker\left(\left[\frac{\partial \phi_1}{\partial x}, \frac{\partial \phi_2}{\partial x}, ..., \frac{\partial \phi_m}{\partial x}\right]^T\right)$$

Ανακεφαλαιώνουμε: Ορίζουμε τον πίνακα

$$d\phi = \left[\frac{\partial \phi_1}{\partial x}, \frac{\partial \phi_2}{\partial x}, ..., \frac{\partial \phi_m}{\partial x}\right]$$

Αν $\delta \in Ker(d\phi^T)$ τότε $d\phi^T \cdot \delta = 0$. Αλλά ταυτόχρονα, $\left(\frac{d\psi}{dx}\right)^T \delta = 0$ στο $x_0$ (αλλιώς το $x_0$ δεν είναι τοπικό βέλτιστο). Άρα, το $\frac{d\psi}{dx}$ είναι γραμμικός συνδιασμός των διανυσμάτων που είναι κάθετα στο $Ker(d\phi)$. Τα διανύσματα αυτά είναι οι στήλες του $d\phi$, οπότε:

$$\frac{d\psi}{dx} = \sum \lambda_i \frac{\partial \phi_i}{\partial x}$$

□

**Παρατήρηση 1**: Με βάση το παραπάνω θεώρημα, αν όλοι οι περιορισμοί είναι γραμμικοί, τότε ο πίνακας των κλίσεών τους, $d\phi$, θα είναι σταθερός, με σταθερό βαθμό σε όλο το χώρο. Άρα είμαστε σίγουροι ότι η συνθήκη $\partial L/\partial x = 0$ είναι αναγκαία και τη χρησιμοποιούμε χωρίς περαιτέρω ελέγχους.

Αν όμως (στην περίπτωση που κάποιος/οι από τους περιορισμούς είναι μη-γραμμικός/οί) κάποια από τα στοιχεία του $d\phi$ εξαρτώνται από το $x$, χρειάζεται περισσότερη ανάλυση για να α) εντοπίσουμε το σύνολο των σημείων στα οποία ο βαθμός του $d\phi$ μεταβάλεται και β) να ελέγξουμε για τυχόν βέλτιστα σε αυτό το σύνολο.

**Παρατήρηση 2**: Η συνθήκη $\partial L/\partial x = 0$ αποτελεί σύστημα $n$ εξισώσεων με $n+m$ αγνώστους (τα $n$ στοιχεία του $x$ και οι $m$ πολλαπλασιαστές Lagrange). Αυτές οι $n$ εξισώσεις μαζί με τους $m$ περιορισμούς $\phi_i(x) = 0$ δίνουν πάντοτε ένα σύστημα $n+m$ εξισώσεων με $n+m$ αγνώστους το οποίο καλούμαστε να λύσουμε για να εντοπίσουμε τα στάσιμα σημεία του προβλήματος.

### 3.3.2 Ικανές συνθήκες για βέλτιστο

Προκειμένου να διερευνήσουμε τις ικανές συνθήκες για την ύπαρξη βέλτιστου σημείου συναρτήσεως με περιορισμούς ισότητας, έστω ότι

$$L = f(x) + \lambda\phi(x) \tag{3.6}$$

όπου $f, \phi : \mathbb{R}^n \to \mathbb{R}$, και $\lambda \in \mathbb{R}$. Υποθέτουμε ότι το $x^*$ είναι τοπικό ελάχιστο για την $f$ και ότι $\phi(x^*) = 0$, ενώ η αντίστοιχη τιμή του $\lambda$ είναι $\lambda^*$. Τότε, για "μικρές" αλλαγές $\delta x, \delta\lambda$ στα $x^*$ και $\lambda$, και προκειμένου το $(x^*, \lambda^*)$ να είναι ελάχιστο, πρέπει να έχουμε

$$\Delta L = L(x^* + \delta x, \lambda^* + \delta\lambda) - L(x^*, \lambda^*) \geq 0$$

Για ευκολία, ορίζουμε το διάνυσμα

$$q = \begin{bmatrix} x \\ \lambda \end{bmatrix}, \quad \delta q = \begin{bmatrix} \delta x \\ \delta\lambda \end{bmatrix}$$

Από την σειρά Taylor της $L$ προκύπτει (όμοια με την προηγούμενη ανάλυση που κάναμε για μονομεταβλητές και πολυμεταβλητές συνάρτησεις χωρίς περιορισμούς) ότι

$$\Delta L = \left( \left.\frac{\partial L}{\partial q}\right|_{x^*,\lambda^*} \right)^T \delta q + \frac{1}{2}\delta q^T \left( \left.\frac{\partial^2 L}{\partial q \partial q}\right|_{x^*,\lambda^*} \right)^T + \ldots \tag{3.7}$$

$$\Delta L = \left( \left.\frac{\partial L}{\partial q}\right|_{x^*,\lambda^*} \right)^T \begin{bmatrix} \delta x \\ \delta\lambda \end{bmatrix} + \frac{1}{2}[\delta x^T \;\; \delta\lambda^T] \left( \left.\frac{\partial^2 L}{\partial q \partial q}\right|_{x^*,\lambda^*} \right)^T \begin{bmatrix} \delta x \\ \delta\lambda \end{bmatrix} + \ldots \tag{3.8}$$

$$\approx \left[\left(\frac{\partial L}{\partial x}\right)^T \ \left(\frac{\partial L}{\partial \lambda}\right)^T\right]\begin{bmatrix}\delta x\\ \delta\lambda\end{bmatrix} + \frac{1}{2}\left(\delta x^T\frac{\partial^2 L}{\partial x\partial x}\delta x + 2\delta x^T\frac{\partial^2 L}{\partial x\partial \lambda}\delta\lambda + \delta\lambda^T\frac{\partial^2 L}{\delta\lambda\delta\lambda}\delta\lambda\right) \quad (3.9)$$

$$= \left(\frac{\partial L}{\partial x}\right)^T\delta x + \left(\frac{\partial L}{\partial \lambda}\right)^T\delta\lambda + \frac{1}{2}\left(\delta x^T\frac{\partial^2 L}{\partial x\partial x}\delta x + 2\delta x^T\frac{\partial^2 L}{\partial x\partial \lambda}\delta\lambda + \delta\lambda^T\frac{\partial^2 L}{\delta\lambda\delta\lambda}\delta\lambda\right) \quad (3.10)$$

$$= \left.\frac{\partial f}{\partial x}\right|_{x^*}\delta x + \lambda^*\left.\frac{\partial \phi}{\partial x}\right|_{x^*}\delta x + \phi(x)\delta\lambda + \frac{1}{2}\left(\delta x^T L_{xx}\delta x + 2\delta x^T L_{x\lambda}\delta\lambda + \delta\lambda^T L_{\lambda\lambda}\delta\lambda\right) \quad (3.11)$$

όπου για συντομία γράψαμε $L_{xx} \triangleq \frac{\partial^2 L}{\partial x\partial x}$, $L_{x\lambda} \triangleq \frac{\partial^2 L}{\partial x\partial \lambda}$, κ.ο.κ.

Στην τελευταία εξίσωση, οι πρώτοι 2 όροι έχουν άθροισμα μηδέν, όπως προκύπτει από την αναγκαία συνθήκη για βέλτιστο στο $x^*$. Επίσης, ο 3ος και ο 6ος όρος είναι ίσοι με μηδέν διότι το $x^*$ ικανοποιεί τους περιορισμούς $\phi(x) = 0$ ενώ η δεύτερη παράγωγος της $L$ ως προς $\lambda$ είναι μηδέν. Από τα παραπάνω, προκύπτει ότι στο σημείο $x^*$:

$$\Delta L = \frac{1}{2}\left(\delta x^T L_{xx}\delta x + 2\delta x^T L_{\lambda x}\delta\lambda\right) + ...$$

άρα το πρόσημο του $\Delta L$ στο τοπικό βέλτιστο ορίζεται από το πρόσημο του όρου

$$[\delta\lambda \ \ \delta x^T]\begin{bmatrix}0 & L_{\lambda x}^T\\ L_{\lambda x} & L_{xx}\end{bmatrix}\begin{bmatrix}\delta\lambda\\ \delta x\end{bmatrix}$$

ή

$$[\delta\lambda \ \ \delta x^T]\hat{H}\begin{bmatrix}\delta\lambda\\ \delta x\end{bmatrix}$$

όπου $\hat{H}$ είναι ο *επαυξημένος* ή *περιφραγμένος* Εσσιανός πίνακας. Άρα, ικανή συνθήκη για να έχουμε τοπικό ελάχιστο είναι η παραπάνω ποσότητα να είναι θετική (ενώ για μέγιστο αρνητική).

*Προσοχή όμως*: Σε αντίθεση με την περίπτωση της βελτιστοποίησης συνάρτησης *χωρίς* περιορισμούς, εδώ η παραπάνω συνθήκη ΔΕΝ μας οδηγεί στο συμπέρασμα ότι $\hat{H} > 0$ (θετικά ορισμένος) για ελάχιστο. Και αυτό γιατί λόγω των περιορισμών, δεν είναι όλα τα $\delta x$ εφικτά, αλλά μόνο αυτά για τα οποία $\phi(x^* + \delta x) = 0$. Για το λόγο αυτό, η συνθήκη $\hat{H} > 0$ είναι υπερβολικά αυστηρή. Η "σωστή" ικανή συνθήκη για τοπικό βέλτιστο σε πρόβλημα με $x \in \mathbb{R}^n$ και $m$ περιορισμούς[1] ισότητας, αφορά μερικές μόνο από τις ηγετικές κύριες ελάσσονες του

$$\hat{H} = \begin{bmatrix}0_{m\times m} & L_{\lambda x}^T\\ L_{\lambda x} & L_{xx}\end{bmatrix}$$

και συγκεκριμένα τις τελευταίες $n - m$ από αυτές.

[1] Ως εδώ υποθέσαμε ότι το πρόβλημά μας είχε μόνο έναν περιορισμό. Στην περίπτωση $m > 1$ περιορισμών, η ανάλυση που ξεκίνησε από την εξίσωση 3.3.2 θα πρέπει να ξαναγίνει, αυτή τη φορά με $L = f(x) + \lambda^T\phi(x)$, όπου $\lambda$ είναι πλέον διάνυσμα πολλαπλασιαστών Lagrange, και $\phi(x) = [\phi_1(x), \phi_2(x), ..., \phi_m(x)]^T$ είναι το διάνυσμα του οποίου τα στοιχεία είναι οι $m$ περιορισμοί ισότητας.

Ο $\hat{H}$ θα έχει συνολικά $(n+m)\times(n+m)$ ηγετικές κύριες ελάσσονες, $\hat{H}_1$,$\hat{H}_2$,...,$\hat{H}_{n+m} = \hat{H}$, όπου $\hat{H}_i$ είναι η $i$ ηγετική κύρια ελάσσονα του $\hat{H}$, δηλαδή ο πίνακας που αποτελείται απο τις πρώτες $i$ στήλες και σειρές του $\hat{H}$. Από τις παραπάνω ελάσσονες, οι πρώτες $m$ είναι μηδενικοί πίνακες, ενώ για $m > 1$ οι $\hat{H}_{m+1}, ..., \hat{H}_{2m-1}$ θα έχουν μηδενική ορίζουσα. Επίσης, μπορεί να δειχθεί εύκολα ότι και η $\hat{H}_{2m}$ *δεν εξαρτάται* από την προς-βελτιστοποίηση συνάρτηση $f$. Μόνο οι τελευταίες (μεγαλύτερες) $n-m$ ελάσσονες, $\hat{H}_{2m+1}, ..., \hat{H}_{k+n} = \hat{H}_n$, παρέχουν πληροφορίες και για την $f$ και για τους περιορισμούς.

Με βάση τα παραπάνω καταλήγουμε στις παρακάτω ικανές συνθήκες:

**Θεώρημα 3.2**

Έστω $x^*$, $\lambda^*$ στάσιμο σημείο (μηδενίζει την παράγωγο)της $L = f(x) + \sum_1^m \lambda_i \phi_i(x)$, και το $x^*$ ικανοποιεί τους περιορισμούς $\phi_i(x) = 0$, $i = 1, ..., m$. Τότε,

- Το $x^*$ είναι τοπικό μέγιστο αν οι τελευταίες $n-m$ ηγετικές κύριες ελάσσονες $\hat{H}_{2m+1}, ..., \hat{H}$ (υπολογισμένες στο σημείο $(x^*, \lambda^*)$) έχουν πρόσημα που εναλλάσονται, ξεκινώντας από το $(-1)^{m+1}$.
- Το $x^*$ είναι τοπικό ελάχιστο αν οι τελευταίες $n-m$ ηγετικές κύριες ελάσσονες $\hat{H}_{2m+1}, ..., \hat{H}$ (υπολογισμένες στο σημείο $(x^*, \lambda^*)$) έχουν το (ίδιο) πρόσημο $(-1)^m$.

**Παράδειγμα 1**

Για να ελαχιστοποιήσουμε την $f(x) = x_1^2 + x_2^2$ υ.π. $\phi(x) = x_1 + x_2 - 1 = 0$ σχηματίζουμε την

$$L = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 1) \tag{3.12}$$

Η αναγκαία συνθήκη για τ.ελάχιστο είναι

$$\frac{\partial L}{\partial x} = \begin{bmatrix} 2x_1 + \lambda \\ 2x_2 + \lambda \end{bmatrix} = 0 \tag{3.13}$$

και φυσικά πρέπει να ισχύει ταυτόχρονα και ο περιορισμός $x_1 + x_2 - 1 = 0$. Η αναγκαία συνθήκη μαζί με τον περιορισμό μας δίνουν ένα σύστημα τριών εξισώσεων με τρεις αγνώστους $x_1, x_2, \lambda$ του οποίου η λύση εύκολα φαίνεται ότι είναι $x_1 = x_2 = 1/2$, $\lambda = -1$.

Για να επιβεβαιώσουμε ότι όντως έχουμε βρει ελάχιστο, αρχικά υπολογίζουμε τον Εσσιανό πίνακα για την $L$,

$$H \triangleq L_{xx} = \begin{bmatrix} \frac{\partial L}{\partial x_1 \partial x_1} & \frac{\partial L}{\partial x_1 \partial x_2} \\ \frac{\partial L}{\partial x_2 \partial x_1} & \frac{\partial L}{\partial x_2 \partial x_2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

οπότε ο επαυξημένος Εσσιανός πίνακας είναι

$$\hat{H} = \begin{bmatrix} 0 & L_{\lambda x}^T \\ L_{\lambda x} & L_{xx} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix} \tag{3.14}$$

Για την κατηγοριοποίηση του στάσιμου σημείου ως μέγιστο ή ελάχιστο μας ενδιαφέρει η $n - m = 2 - 1 = 1$ τελευταία ηγετική κύρια ελάσσονα - δηλαδή ολόκληρος ο $\hat{H}$. Η ορίζουσά του ισούται με $-4 < 0$, και το πρόσημό της είναι ίδιο με αυτό του $(-1)^m = -1$ (ο $m = 1$ είναι περιττός - ταιριάζει με το "μοτίβο" -,-,-,..-, για ελάχιστο). Άρα το $x = [1/2, 1/2]^T$ αποτελεί ελάχιστο.

**Παράδειγμα 2**

Για να μεγιστοποιήσουμε την $f(x) = x_1^2 + x_2^2$ υ.π. $5x_1^2 + 6x_1x_2 + 5x_2^2 = 8$ σχηματίζουμε την

$$L = x_1^2 + x_2^2 + \lambda(5x_1^2 + 6x_1x_2 + 5x_2^2 - 8) \tag{3.15}$$

Η αναγκαία συνθήκη για τ.μέγιστο είναι

$$\frac{\partial L}{\partial x} = \begin{bmatrix} 2x_1 + 10x_1\lambda + 6\lambda x_2 \\ 2x_2 + 6\lambda x_1 + 10\lambda x_2 \end{bmatrix} = 0 \tag{3.16}$$

Από την πρώτη εκ των δύο παραπάνω εξισώσεων, βλέπουμε ότι

$$\lambda = \frac{-2x_1}{10x_1 + 6x_2} \tag{3.17}$$

οπότε (από την δεύτερη εξίσωση της 3.16)

$$2x_2 + \frac{-2x_1}{10x_1 + 6x_2}6x_1 + 10x_2\frac{-2x_1}{10x_1 + 6x_2} = 0 \Rightarrow \tag{3.18}$$

$$20x_1x_2 + 12x_2^2 - 12x_1^2 - 20x_1x_2 = 0 \Rightarrow \tag{3.19}$$

$$x_2^2 - x_1^2 = 0 \Rightarrow x_1 = \pm x_2 \tag{3.20}$$

και άρα πρέπει να εξετάσουμε τις 2 περιπτώσεις $x_1 = x_2$, και $x_1 = -x_2$ χωριστά.

Εάν $x_1 = x_2$, τότε η (3.17) δίνει $\lambda = \frac{-2x_1}{16x_1}$, και από την εξίσωση του περιορισμού έχουμε

$$5x_1^2 + 6x_1^2 + 5x_1^2 = 8 \Rightarrow x_1 = \pm 1/\sqrt{2}$$

Για τις τιμές αυτές των $x_1, x_2$, έχουμε $\lambda = -1/8$. Η τιμή της αντικειμενικής συνάρτησης $f$ για $x_1 = x_2 = \pm 1/\sqrt{2}$ είναι $f = 1$. Για να βρούμε αν η τιμή $f = 1$ είναι μέγιστη (αντί για ελάχιστη), θα μπορούσαμε να εξετάσουμε την τιμή της $f$ στην δεύτερη περίπτωση (όπου $x_1 = -x_2$, όπου ανάλογη σειρά υπολογισμών μας οδηγεί στο $x_1 = \pm\sqrt{2}$, $x_1 = \mp\sqrt{2}$, $\lambda = -1/2$, $f = 4$ - οπότε φαίνεται ότι σε αυτό το σημείο βρίσκεται το μέγιστο ενώ στο προηγούμενο έχουμε ελάχιστο), αλλά και τις κατάλληλες ελάσσονες του επαυξημένου Εσσιανού πίνακα.

Αρχικά υπολογίζουμε τον Εσσιανό πίνακα για την $L$,

$$H \triangleq L_{xx} = \begin{bmatrix} \frac{\partial L}{\partial x_1 \partial x_1} & \frac{\partial L}{\partial x_1 \partial x_2} \\ \frac{\partial L}{\partial x_2 \partial x_1} & \frac{\partial L}{\partial x_2 \partial x_2} \end{bmatrix} = \begin{bmatrix} 2 + 10\lambda & 6\lambda \\ 6\lambda & 2 + 10\lambda \end{bmatrix}$$

οπότε ο επαυξημένος Εσσιανός πίνακας είναι

$$\hat{H} = \begin{bmatrix} 0 & L_{\lambda x}^T \\ L_{\lambda x} & L_{xx} \end{bmatrix} \tag{3.21}$$

$$= \left[\begin{array}{c|c|c} 0 & 10x_1 + 6x_2 & 6x_1 + 10x_2 \\ \hline 10x_1 + 6x_2 & 2 + 10\lambda & 6\lambda \\ 6x_1 + 10x_2 & 6\lambda & 2 + 10\lambda \end{array}\right] \tag{3.22}$$

Επιστρέφοντας στο ζεύγος στάσιμων σημείων $x = (\pm\sqrt{2}, \mp\sqrt{2})$, $\lambda = -1/2$ το οποίο εντοπίσαμε ως πιθανό μέγιστο, αντικαταθιστούμε στον επαυξημένο Εσσιανό πίνακα ( έστω για $x = (\sqrt{2}, -\sqrt{2})$ - η ανάλυση για το $x = (-\sqrt{2}, \sqrt{2})$ είναι όμοια) και έχουμε:

$$\hat{H} = \begin{bmatrix} 0 & 10\sqrt{2} - 6\sqrt{2} & 6\sqrt{2} - 10\sqrt{2} \\ 10\sqrt{2} - 6\sqrt{2} & -3 & -3 \\ 6\sqrt{2} - 10\sqrt{2} & -3 & -3 \end{bmatrix}$$

ή

$$\hat{H} = \begin{bmatrix} 0 & 4\sqrt{2} & -4\sqrt{2} \\ 4\sqrt{2} & -3 & -3 \\ -4\sqrt{2} & -3 & -3 \end{bmatrix}$$

Για την κατηγοριοποίηση του στάσιμου σημείου ως μέγιστο ή ελάχιστο μας ενδιαφέρει η $n - m = 2 - 1 = 1$ τελευταία ηγετική κύρια ελάσσονα - δηλαδή ολόκληρος ο $\hat{H}$. Η ορίζουσά του ισούται με $384 > 0$, και το πρόσημό της είναι ίδιο με αυτό του $(-1)^{m+1} = 1$ (ο $m$ είναι περιττός - ταιριάζει με το "μοτίβο" +,-,+,.. για μέγιστο). Άρα το $x = (\sqrt{2}, -\sqrt{2})$ (αλλά και το $x = (-\sqrt{2}, \sqrt{2})$ αν κάνουμε τις αντίστοιχες πράξεις) αποτελεί μέγιστο.

**Παράδειγμα 3**

Έστω ότι θέλουμε να ελαχιστοποιήσουμε την $f(x) = x_1^2 + 2x_2^2 + (x_3 - 2)^2 + x_1 x_2 + 2x_1 x_3$ υ.π. $x_1 + x_2 + x_3 = 1$ και $2x_1 + x_2 - x_3 = 5$.

Πρόκειται για πρόβλημα μη-γραμμικού προγραμματισμού με $n = 3$ μεταβλητές και $m = 2$ περιορισμούς. Ξεκινάμε γράφοντας τους δύο περιορισμούς στην τυποποιημένη μορφή, $\phi_1(x) = x_1 + x_2 + x_3 - 1$ και $\phi_2(x) = 2x_1 + x_2 - x_3 - 5 = 0$, και σχηματίζουμε τη Λαγκραντζιανή:

$$L = x_1^2 + 2x_2^2 + (x_3 - 2)^2 + x_1 x_2 + 2x_1 x_3 + \lambda_1(x_1 + x_2 + x_3 - 1) + \lambda_2(2x_1 + x_2 - x_3 - 5) \tag{3.23}$$

Η αναγκαία συνθήκη για βέλτιστο είναι:

$$\frac{\partial L}{\partial x} = \begin{bmatrix} 2x_1 + x_2 + 2x_3 + \lambda_1 + 2\lambda_2 \\ 4x_2 + x_1 + \lambda_1 + \lambda_2 \\ 2x_3 - 4 + 2x_1 + 2\lambda_1 - \lambda_2 \end{bmatrix} = 0 \tag{3.24}$$

Οι 3 αυτές εξισώσεις, μαζί με τις εξισώσεις $\phi_1(x) = 0$, $\phi_2(x) = 0$ μας δίνουν ένα $5 \times 5$ σύστημα το οποίο πρέπει να λυθεί. Παρατηρούμε ότι και οι πέντε εξισώσεις συμβαίνει να είναι γραμμικές. Οπότε μπορούμε να γράψουμε το σύστημα στη μορφή:

$$Az = b, \text{όπου } A = \begin{bmatrix} 2 & 1 & 2 & 1 & 2 \\ 1 & 4 & 0 & 1 & 1 \\ 2 & 0 & 2 & 2 & -1 \\ 1 & 1 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 0 \end{bmatrix}, \quad z = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda_1 \\ \lambda_2 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ 4 \\ 1 \\ 5 \end{bmatrix}.$$

Λύνοντας το σύστημα (είτε με απαλοιφή Gauss είτε αντιστρέφοντας τον πίνακα $A$) παίρνουμε

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 2.185 \\ -0.277 \\ -0.908 \\ 0.123 \\ -1.200 \end{bmatrix}$$

Το οποίο είναι το μοναδικό πιθανό ελάχιστο. Για να διαπιστώσουμε αν όντως είναι ελάχιστο, θα πρέπει να σχηματίσουμε τον επαυξημένο Εσσιανό πίνακα ο οποίος θα έχει διαστάσεις $5 \times 5$ (γιατί $n + m = 5$) και να ελέγξουμε την τελευταία $n - m = 3 - 2 = 1$ ορίζουσά του (δηλ. την $5 \times 5$ ορίζουσα ολόκληρου του πίνακα). Έχουμε λοιπόν:

$$\hat{H} = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & -1 \\ \hline 1 & 2 & 2 & 1 & 2 \\ 1 & 1 & 1 & 4 & 0 \\ 1 & -1 & 2 & 0 & 2 \end{array}\right]$$

Η ορίζουσα του $\hat{H}$ είναι $det(\hat{H}) = 42 > 0$. Επειδή το πρόβλημά μας έχει άρτιο αριθμό περιορισμών, αυτό αντιστοιχεί στο μοτίβο προσήμων "+ + +...+" το οποίο δηλώνει ότι το σημείο που υπολογίσαμε είναι ελάχιστο.

Στα προηγούμενα παραδείγματα, ο στόχος μας ήταν είτε α) η βελτιστοποίηση δευτεροβάθμιας αντικειμενικής συνάρτησης με γραμμικούς περιορισμούς, είτε β) βελτιστοποίηση δευτεροβάθμιας συνάρτησης με δευτεροβάθμιο περιορισμό. Τα επόμενα δύο παραδείγματα δείχνουν πως μπορούμε να επιλύσουμε τις δύο αυτές κατηγορίες προβλημάτων στη γενική τους μορφή, ώστε να καταλήξουμε σε λύσεις/τύπους οι οποίοι δεν εξαρτώνται από το πλήθος των μεταβλητών και δεν απαιτούν να χειριζόμαστε την κάθε μεταβλητή $x_1, x_2, ..., x_n$ χωριστά.

**Παράδειγμα 4**

Έστω ότι θέλουμε να ελαχιστοποιήσουμε την τετραγωνική μορφή (quadratic form)

$$f(x) = x^T Q x$$

υ.π

$$C^T x = b, \text{όπου} x, C \in \mathbb{R}^n, \ Q = Q^T > 0.$$

Παρατηρούμε ότι $x^T Q x \in \mathbb{R}$ (δηλαδή η $f(x)$ είναι πραγματικός για κάθε $x$) και, δεδομένου ότι ο $Q$ είναι θετικά ορισμένος, θα ισχύει $f(x) = x^T Q x > 0$ για $x \neq 0$ (βλ. Παράρτημα Α').

Στο $\mathbb{R}^2$, το πρόβλημα αυτό είναι ισοδύναμο με την εύρεση της μικρότερης έλλειψης η οποία τέμνει την ευθεία $C^T x = b$ (*Άσκηση: να γίνει ενδεικτική γραφική παράσταση)*.

Παρατηρούμε ότι ο πίνακας $d\phi = C$ έχει σταθερή τάξη (διότι ο $C$ είναι σταθερός), άρα είμαστε βέβαιοι ότι η αναγκαία συνθήκη $\partial L/\partial x = 0$ θα ισχύει στο σημείο που ελαχιστοποιείται η $f$. Οπότε προχωράμε σχηματίζοντας την Λαγκραντζιανή:

$$L = x^T Q x + \lambda(C^T x - b)$$

και υπολογίζουμε

$$\partial L/\partial x = 2Qx + C\lambda = 0$$

ή:

$$2Qx = -C\lambda, \tag{3.25}$$

το οποίο πρέπει να ισχύει στο τοπικό ελάχιστο. Παρατηρούμε ότι και εδώ, η τελευταία συνθήκη απαιτεί **η παράγωγος της** $f$ (αριστερή πλευρά) **να είναι γραμμικός συνδυασμός των παραγώγων των περιορισμών** (δεξιά πλευρά). Με άλλα λόγια, αν το $x_0$ είναι τοπικό ελάχιστο τότε θα πρέπει η α) κάθετη στην ισοσταθμική επιφάνεια $f(x_0) =$ σταθερά και η β) κάθετη στην επιφάνεια που ορίζεται από τους περιορισμούς, να είναι παράλληλες στο σημείο $x_0$.

Επειδή $Q = Q^T > 0$ γνωρίζουμε ότι ο $Q$ θα είναι αντιστρέψιμος[2], έχουμε

$$2Qx = -C\lambda \Rightarrow x = -\frac{1}{2}Q^{-1}C\lambda$$

Βλέπουμε ότι χρειάζεται να υπολογίσουμε τον άγνωστο πολλαπλασιαστή Lagrange, $\lambda$. Αυτό γίνεται εύκολα χρησιμοποιόντας τον περιορισμό που δόθηκε με το πρόβλημα:

$$C^T x - b = 0 \Leftrightarrow -\frac{1}{2}C^T Q^{-1} C\lambda = b \Leftrightarrow \lambda = (CQ^{-1}C)^{-1}(-2b)$$

και τελικά,

$$x = Q^{-1}C(CQ^{-1}C)^{-1}b$$

Η ελάχιστη τιμή της $f$ είναι

$$f_{min} = x^T Q x = b^T (CQ^{-1}C)^{-1} C Q^{-1} C (CQ^{-1}C)^{-1} b$$

**Σημείωση:** Η παραπάνω έκφραση μπορεί να απλοποιηθεί σε $f_{min} = b^2/(C^T Q^{-1} C)$ στην περίπτωση που έχουμε μόνο έναν περιορισμό (αντίστοιχα ο $C$ αποτελείται από μία μόνο στήλη). Στην περίπτωση πολλών περιορισμών, ο τύπος για την ελάχιστη τιμή της $f$ δεν μπορεί να απλοποιηθεί περαιτέρω και η προηγούμενη ανάλυση είναι σωστή αρκεί να θεωρήσουμε ότι η ποσότητα $\lambda$ είναι πλέον διάνυσμα (με ένα στοιχείο για κάθε περιορισμό) και να αντικαταστήσουμε το $\lambda$ με $\lambda^T$ στη Λαγκραντζιανή που γράψαμε αρχικά - στις υπόλοιπες εξισώσεις δε χρειάζεται καμμια αλλαγή.

[2] Οι θετικά ορισμένοι πίνακες έχουν όλες τις ιδιοτιμές τους αυστηρά θετικές. Η ορίζουσα του $Q$ θα είναι - κατά συνέπεια - αυστηρά θετική (επειδή η ορίζουσα ισούται με το γινόμενο των ιδιοτιμών), άρα μη-μηδενική, οπότε υπάρχει ο $Q^{-1}$.

**Παράδειγμα 5**

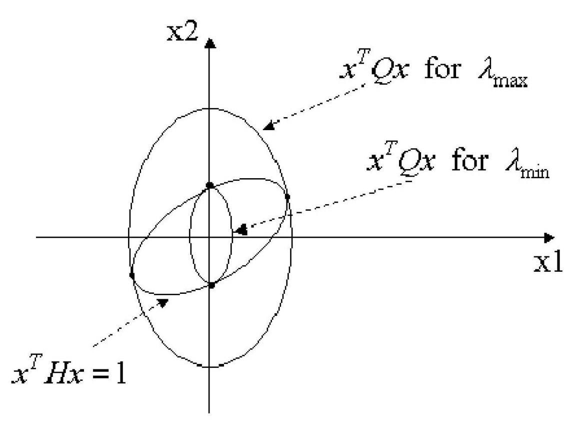
$$Min\, x^T Qx \quad \text{υ.π.} \quad x^T Hx = 1,\ \text{όπου}\ Q = Q^T > 0,\ H = H^T > 0$$

Σχηματίζουμε τη Λαγκραντζιανή:

$$\begin{aligned} L = x^T Qx + \lambda(x^T Hx - 1) \Rightarrow \\ \frac{\partial L}{\partial x} = Qx + \lambda Hx = 0 \Rightarrow \\ (Q + \lambda H)x = 0 \Rightarrow \\ det(Q + \lambda H) = 0 \Rightarrow \\ det(QH^{-1} + \lambda I)x = 0 \end{aligned}$$

το οποίο σημαίνει ότι ο $\lambda$ **είναι ιδιοτιμή του πίνακα** $-QH^{-1}$ (σημειώνουμε ότι οι θετικά ορισμένοι πίνακες έχουν αυστηρά θετικές ιδιοτιμές, οπότε οι αντίστροφοι που εμφανίζονται στις παραπάνω σχέσεις είναι σίγουρο οτι υπάρχουν).

Αν οι $Q, H$ είναι $n \times n$, τότε υπάρχουν $n$ πιθανές λύσεις για το $\lambda$. Ποιά είναι αυτή που αντιστοιχεί στο ελάχιστο; Από το διάγραμμα του προβλήματος (Σχ.3.1) (που αντιπροσωπεύει μια




**Σχήμα 3.1.** Τα σημεία τομής των $x^T Hx = 1$ και $x^T Qx =$ σταθ..

έκδοση του προβλήματος σε 2 διαστάσεις) βλέπουμε ότι πρέπει να επιλέξουμε την "εσωτερική" έλλειψη για το ελάχιστο (ή την "εξωτερική" για το μέγιστο). Γενικά θα υπάρχουν $2n$ επιλογές για τα $x$ που ικανοποιούν τις αναγκαίες συνθήκες για ελάχιστο ή μέγιστο (υπάρχουν $n$ ιδιοτιμές ενώ τα $-x, x$ δίνουν και τα δύο την ίδια τιμή για το $x^T Qx$). Για να βρούμε αυτές τις $(2n)$ υποψήφιες λύσεις, ανατρέχουμε στις αναγκαίες συνθήκες για βέλτιστο

$$Qx + \lambda Hx = 0 \Rightarrow \tag{3.26}$$

$$QH^{-1}Hx = -\lambda Hx \Rightarrow \tag{3.27}$$

$$-QH^{-1}y = \lambda y \tag{3.28}$$

όπου έχουμε ορίσει $y \stackrel{\triangle}{=} Hx$. Από την τελευταία εξίσωση, βλέπουμε ότι το $y$ πρέπει να είναι ιδιοδιάνυσμα του $-QH^{-1}$. Υπάρχουν $n$ τέτοια ιδιοδιανύσματα τα οποία συμβολίζουμε με $y_i, i = 1, ..., n$. Οι αντίστοιχες λύσεις για το $x$ θα είναι $x_i = H^{-1}y_i$.

Παρατηρούμε (από την (3.26)) ότι τα $x_i$ μπορούν να πολλαπλασιαστούν με οποιοδήποτε πραγματικό χωρίς να παραβιάσουν τις αναγκαίες συνθήκες για βέλτιστο (αν το $x_i$ είναι ιδιοδιάνυσμα ενός πίνακα, τότε το ίδιο ισχύει και για όλα τα πολλαπλάσια του $x_i$). Πολλαπλασιάζουμε λοιπόν το κάθε $x_i$ με τον κατάλληλο παράγοντα ώστε τα ικανοποιεί τον περιορισμό $x_i^T Hx_i = 1$. Τα νέα αυτά $x_i$, μαζί με τα $-x_i$, θα είναι οι $2n$ πιθανές λύσεις στο πρόβλημα ελαχιστοποίησης του $x^T Qx$.

Στο σημείο αυτό, θα μπορούσαμε απλώς να υπολογίσουμε την αντικειμενική συνάρτηση σε κάθε ένα από τα $2n$ σημεία που αναφέραμε, και να επιλέξουμε έτσι το ελάχιστο. Υπάρχει όμως καλύτερος τρόπος για να λύσουμε το πρόβλημα:

Εξετάζοντας τον πίνακα των δευτέρων παραγώγων της $L$, $\frac{\partial^2 L}{\partial x \partial x} = Q + \lambda H$, βλέπουμε ότι είναι συμμετρικός(διότι και ο $Q$ και ο $H$ είναι συμμετρικοί). Επειδή όπως δείξαμε ήδη, το $\lambda$ πρέπει να είναι μία από τις ιδιοτιμές του $-QH^{-1}$, θα πρέπει να ισχύει $det(Q+\lambda H) = 0$. Τώρα, ισχυριζόμαστε ότι όλες οι λύσεις (ως προς $\lambda$) της $det(Q + \lambda H) = 0$ πρέπει να είναι αρνητικές (αν $\lambda \geq 0$, τότε $Q + \lambda H$ είναι συμμετρικός θετικά ορισμένος, οπότε η ορίζουσά του δεν μπορεί να είναι μηδέν). Άρα

$$\partial L/\partial x = 0 \Rightarrow Qx + \lambda Hx = 0 \Rightarrow x^T Qx + x^T Hx\lambda = 0 \Rightarrow x^T Qx = -\lambda > 0 \tag{3.29}$$

όπου $\lambda$ είναι ιδιοτιμή του $-QH^{-1}$. Έχουμε λοιπόν,

$$Min_x x^T Qx = -\lambda_{max} \quad (\text{ λιγότερο αρνητική ιδιοτιμή } \; -QH^{-1})$$

$$Max_x x^T Qx = -\lambda_{min} \quad (\text{ περισσότερο αρνητική ιδιοτιμή } \; -QH^{-1})$$

Για να εμπεδώσουμε τα παραπάνω, ας υπολογίσουμε τα βέλτιστα $x$ και τη μέγιστη και ελάχιστη τιμή της $f$ για συγκεκριμένες τιμές των $Q, H$:

$$Q = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}.$$

Οι δύο αυτοί πίνακες είναι προφανώς συμμετρικοί, και (ελέγχοντας τις ηγετικές κύριες ελάσσονες) θετικά ορισμένοι.

Έχουμε

$$H^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}, \qquad -QH^{-1} = \begin{bmatrix} -3 & -1/3 \\ -1 & -2/3 \end{bmatrix}$$

Από τις ρίζες του πολυωνύμου $det(-QH^{-1} - \lambda I) = 0$ παίρνουμε τις ιδιοτιμές $\lambda_1 = -3.135$, $\lambda_2 = -0.5316$. Από αυτές, γνωρίζουμε ήδη ότι η μέγιστη τιμή της $f$ είναι 3.135 ενώ η ελάχιστη 0.5316.

Συνεχίζουμε για να βρουμε και τα σημεία $x$ στα οποία επιτυγχάνεται το ελάχιστο και το μέγιστο. Για τα ιδιοδιανύσματα $y$ του πίνακα $-QH^{-1}$ λύνουμε το σύστημα

$$-QH^{-1}y = \lambda y$$

για $\lambda = \lambda_1$ και για $\lambda = \lambda_1$ και έχουμε τα αντίστοιχα ιδιοδιανύσματα:

$$y_1 = \begin{bmatrix} -0.9268 \\ -0.3755 \end{bmatrix}, \quad y_2 = \begin{bmatrix} 0.1338 \\ -0.9910 \end{bmatrix}$$

(προσοχή - τα ιδιοδιανύσματα ως γνωστόν δεν είναι μοναδικά, οπότε οποιαδήποτε πολλαπλάσια των παραπάνω είναι το ίδιο “καλά”).

Στη συνέχεια υπολογίζουμε τα $x_1 = H^{-1}y_1$ και $x_2 = H^{-1}y_2$:

$$x_1 = \begin{bmatrix} -0.9268 \\ -0.1252 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0.1338 \\ -0.3303 \end{bmatrix}$$

Δοκιμάζουμε αν αυτές οι λύσεις ικανοποιούν τον περιορισμό $x^THx = 1$. Κάνοντας τις πράξεις, βρίσκουμε ότι

$$x_1^THx_1 = 0.9060, \quad x_2^THx_2 = 0.3453.$$

Αυτό συμβαίνει γιατί αν και οι διευθύνσεις των $y_1, y_2$ που υπολογίσαμε είναι σωστές, τα μέτρα τους δεν ήταν (θυμίζουμε ότι τα ιδιοδιανύσματα ορίζονται με μοναδικό τρόπο μόνο ως προς τη διεύθυνσή τους). Θα χρειαστεί λοιπόν να διαιρέσουμε τις λύσεις μας με τις κατάλληλες σταθερές ώστε να δίνουν πάντα $x^THx = 1$. Οι σταθερές αυτές είναι οι τετραγωνικές ρίζες των παραπάνω αποτελεσμάτων, $\sqrt{0.9060}, \sqrt{0.3453}$. Διαιρώντας με αυτές τις τιμές τα $x_1$ και $x_2$ αντίστοιχα, έχουμε

$$x_1^* = \begin{bmatrix} -0.9737 \\ -0.1315 \end{bmatrix}, \quad x_2^* = \begin{bmatrix} 0.2277 \\ -0.5622 \end{bmatrix}$$

οι οποίες είναι και οι βέλτιστες λύσεις (η $x_1^*$ για μέγιστο, η $x_2^*$ για το ελάχιστο). Μπορούμε πλέον να επιβεβαιώσουμε κάνοντας τους αντίστοιχους υπολογισμούς ότι π.χ. $f_{min} = x_2^*Qx_2^* = 0.5316$ ενώ $x_2^*Hx_2^* = 1$ όπως απαιτείται.

Παρατήρηση: Η αξία των δύο παραπάνω παραδειγμάτων γίνεται προφανής σε χώρους μεγάλης διάστασης ($n$ μεγάλο), όπου αν δεν εκμεταλλευόμασταν τον τρόπο επίλυσης που ακολουθήσαμε (χειριστήκαμε το διάνυσμα $x$ σαν ένα ενιαίο αντικείμενο και καταλήξαμε σε τύπους για τη βέλτιστη λύση, οι οποίοι απαιτούν απλή αντικατάσταση των παραμέτρων του προβλήματος), θα έπρεπε να σχηματίζουμε και να επιλύουμε μεγάλα συστήματα εξισώσεων, για κάθε νέα επιλογή των παραμέτρων $Q, c, b, H$, κάτι που θα ήταν πιο επίπονο και χρονοβόρο.

### 3.3.3 Εφαρμογή: Γραμμικά συστήματα εξισώσεων

Ας θεωρήσουμε το γραμμικό σύστημα

$$Ax = b, \quad A : m \times n,\ x \in \mathbb{R}^n,\ b \in \mathbb{R}^m$$

Είναι γνωστό ότι στην περίπτωση $n = m$ με $det(A) \neq 0$, το σύστημα έχει μοναδική λύση την $x = A^{-1}b$. Δύο ενδιαφέροντα προβλήματα βελτιστοποίησης προκύπτουν στις περιπτώσεις $m < n$ (περισσότερες μεταβλητές και λιγότεροι περιορισμοί) και $m > n$ (λιγότερες μεταβλητές και περισσότεροι περιορισμοί).

### Περίπτωση $m > n$

Στην περίπτωση αυτή οι βαθμοί ελευθερίας που διαθέτουμε για να λύσουμε το σύστημα είναι λιγότεροι από τις απαιτήσεις/περιορισμούς που θέτουμε σε αυτές, και γενικά το σύστημα δε θα έχει λύση (εκτός από τις περιπτώσεις όπου το $b$ ανήκει στο χώρο στηλών του $A$ ή, αντίστοιχα, κάποιοι περιορισμοί είναι πλεονασματικοί). Θα υποθέσουμε ότι $rank(A) = n$ - δηλαδή όλες οι στήλες του $A$ είναι γραμμικά ανεξάρτητες, και - μιας και το σύστημα $Ax = b$ δεν έχει λύση, θα αναζητήσουμε το $x$ το οποίο είναι "πλησιέστερο" στο να είναι λύση, με την έννοια ότι ελαχιστοποιεί την απόσταση μεταξύ $Ax$ και $b$. Αυτό μας οδηγεί στο παρακάτω πρόβλημα βελστιστοποίησης:

$$Min_x ||Ax - b||$$

όπου $||x|| = \sqrt{x^T x} = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$ είναι το ευκλείδιο μήκος του διανύσματος $x$.

Για να διευκολύνουμε τα πράγματα, παρατηρούμε ότι η ελαχιστοποίηση του μήκους ενός διανύσματος είναι ισοδύναμη με την ελαχιστοποίηση του μήκους του στο τετράγωνο. Άρα, το παραπάνω πρόβλημα είναι ισοδύναμο με το

$$Min_x ||Ax - b||^2 = (Ax - b)^T (Ax - b) = x^T A^T Ax - b^T Ax - x^T A^T b + b^T b$$

η οποία είναι δευτεροβάθμια συνάρτηση της διανυσματικής μεταβλητής $x$. Για να την ελαχιστοποιήσουμε, αναγκαία συνθήκη είναι ο μηδενισμός της παραγώγου της:

$$\frac{\partial}{\partial x}(x^T A^T Ax - b^T Ax - x^T A^T b + b^T b) = 0 \Rightarrow 2A^T Ax - 2A^T b = 0$$

όπου για την παραγώγιση χρησιμοποιήσαμε τους κανόνες για γραμμικές και τετραγωνικές συναρτήσεις διανυσματικής μεταβλητής οι οποίοι περιγράφονται στο Παράρτημα Β. Η τελευταία εξίσωση, μας δίνει

$$x = (A^T A)^{-1} A^T b$$

όπου ο απαραίτητος αντίστροφος στη δεξιά πλευρά θα υπάρχει λόγω της υπόθεσης που κάναμε ότι $rank(A) = n$. Με τον τρόπο αυτό βρήκαμε το διάνυσμα $x$ το οποίο φέρνει την αριστερή πλευρά του συστήματος $Ax = b$ όσο πλησιέστερα γίνεται στη δεξιά.

### Περίπτωση $n > m$

Στην περίπτωση που οι μεταβλητές του συστήματος είναι περισσότερες από τους περιορισμούς, το σύστημα έχει γενικά άπειρες λύσεις (εξαιρώντας "παθολογικές" περιπτώσεις όπου κάποιες από τις εξισώσεις είναι ασύμβατες). Στην περίπτωση αυτή, μπορούμε να αναζητήσουμε την ποιο "οικονομική" από τις λύσεις, ως αυτή με το ελάχιστο άθροισμα τετραγώνων:

$$Min||x||^2 \quad \text{υ.π.} Ax = b.$$

Τώρα έχουμε ένα πρόβλημα ελαχιστοποίησης με περιορισμούς ισότητας, άρα ακολουθώντας τη μέθοδο πολλαπλασιαστών Lagrange,

$$L = x^T x + \lambda(Ax - b)$$

όπου $\lambda \in \mathbb{R}^m$ είναι ένα διάνυσμα πολλαπλασιαστών Lagrange, ένας για κάθε περιορισμό (σειρά του $A$). Η αναγκαία συνθήκη είναι

$$\frac{\partial L}{\partial x} = \frac{\partial}{\partial x}(x^T x + \lambda(Ax - b)) = 2x + A^T\lambda = 0$$

η οποία πρέπει να λυθεί για $x, \lambda$ μαζί με τους περιορισμούς $Ax = b$. Έχουμε

$$x = -\frac{1}{2}A^T\lambda$$

από την οποία η $Ax = b$ μας δίνει

$$-\frac{1}{2}A^T\lambda = b \Rightarrow \lambda = -2(AA^T)^{-1}b$$

Αντικαθιστώντας για το $\lambda$ στον τύπο για το $x$ έχουμε τελικά

$$x = A^T(AA^T)^{-1}b.$$

το οποίο είναι η λύση του γραμμικού συστήματος με το μικρότερο δυνατό μήκος.

## 3.4 Βελτιστοποίηση με Περιορισμούς Ανισότητας

Προτού καταπιαστούμε με προβλήματα βελτιστοποίησης υπό περιορισμούς ανισότητας, είναι χρήσιμο να εξετάσουμε σε ποια σημεία της εφικτής περιοχής αναμένεται να υπάρχει βέλτιστο. Στην περίπτωση που η εφικτή περιοχή είναι κυρτή, η απάντηση είναι απλή, και έχει να κάνει με το αν η αντικειμενική συνάρτηση είναι κοίλη ή κυρτή.

### 3.4.1 Κυρτότητα - Βασικές έννοιες

**Ορισμός 4.1**

Το σύνολο $S$ είναι *κυρτό* άν $\forall\, x_1, x_2 \in S$ και $\forall\, 0 \leq \theta \leq 1$, έχουμε $\theta x_1 + (1-\theta)x_2 \in S$.

Δηλαδή το $S$ είναι κυρτό αν για κάθε 2 σημεία του $S$, το ευθύγραμμο τμήμα που τα ενώνει ανήκει στο $S$.

**Ορισμός 4.2**

Το σημείο $x$ είναι ακρότατο του κυρτού συνόλου $S$ εάν δεν υπάρχουν σημεία $x_1, x_2 \in S$ τέτοια ώστε να είναι

$$x = \theta x_1 + (1-\theta)x_2$$

για $0 < \theta < 1$. Δηλαδή το $x$ δεν ανήκει σε ευθύγραμμο τμήμα του οποίου τα άκρα είναι στο $S$.

**Θεώρημα 4.1**

Έστω $S_1, S_2, ..., S_N$ κυρτά σύνολα. Τότε και η τομή τους $S = \cap_1^N S_i$ είναι κυρτό σύνολο.

Απόδειξη: Έστω $x_1, x_2 \in S$, τότε $x_1, x_2 \in S_i \; \forall \, i$, που σημαίνει ότι όλα τα σημεία

$$x = \theta x_1 + (1-\theta)x_2, \quad \theta \in [0,1]$$

ανήκουν στο $S_i$, για κάθε $i$, άρα ανήκουν και στην τομή των συνόλων $S_i$. $\square$

**Ορισμός 4.3**

Έστω $K$ κυρτό σύνολο και $h(x) : K \to \mathbb{R}$ παραγωγίσιμη συνάρτηση. Λέμε ότι η $h$ είναι κοίλη αν

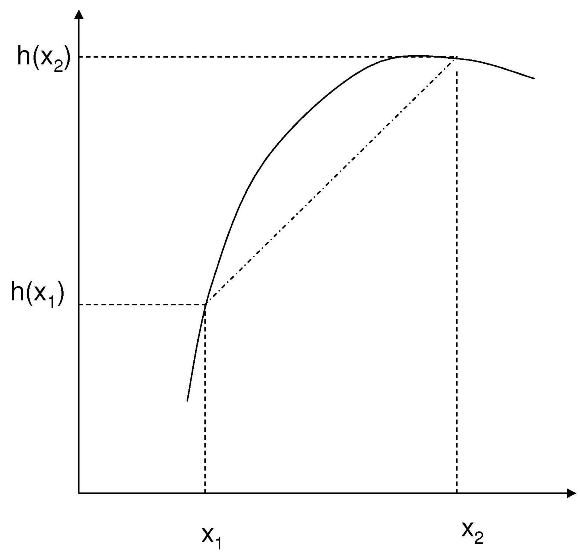
$$\forall \, x_1, x_2 \in K : h(\theta x_1 + (1-\theta)x_2) \geq \theta h(x_1) + (1-\theta)h(x_2)$$

(βλ. Σχ. 3.2).

Για να ορίσουμε την έννοια της *κυρτής* συνάρτησης αντιστρέφουμε τη φορά της ανισότητας στον παραπάνω ορισμό.

**Θεώρημα 4.2**

Έστω $K$ κυρτό σύνολο και $h(x) : K \to \mathbb{R}$ παραγωγίσιμη συνάρτηση. Η $h$ είναι κοίλη αν και μόνο αν $\forall \, x_1, x_2 \in K$:




**Σχήμα 3.2.** Ορισμός κοίλης συνάρτησης στο $\mathbb{R}$. Η $h$ μεταξύ $x_1$ και $x_2$ πρέπει να είναι βρίσκεται "πάνω" από το ευθύγραμμο τμήμα που ενώνει τα $(x_1, h(x_1))$ και $(y_1, h(y_2))$.

$$f(x_2) - f(x_1) \leq \left( \frac{\partial f}{\partial x}\bigg|_{x_1} \right)^T (x_2 - x_1)$$

Από τα παραπάνω συμπεραίνουμε ότι αν η $h$ είναι κυρτή τότε η $-h$ είναι κοίλη, και αντίστροφα, και ότι οι γραμμικές συναρτήσεις είναι κοίλες και κυρτές ταυτόχρονα. Επίσης με βάση τον παραπάνω ορισμό είναι δυνατό να αποδειχθεί ότι το παρακάτω.

**Θεώρημα 4.3**

Έστω ότι η $h : K \to \mathbb{R}$, όπου $K$ κυρτό, είναι κοίλη συνάρτηση. Αν η $h$ έχει ολικό ελάχιστο στο $x^*$, τότε ένα η περισσότερα $x^*$ βρίσκονται στα ακρότατα του $K$.

Φυσικά, αν η $h$ έχει ολικό μέγιστο, αυτό μπορεί να βρίσκεται είτε στα ακρότατα είτε στο εσωτερικό του $K$. Για $h$ κυρτή, τα παραπάνω αντιστρέφονται, δηλαδή το ολικό μέγιστο είναι αυτό που βρίσκεται σε ακρότατο του .

**Θεώρημα 4.4**

Έστω ότι η $h : K \to \mathbb{R}$, όπου $K$ κυρτό, είναι κοίλη συνάρτηση. Για κάθε σταθερά $\gamma$, το σύνολο $H_\gamma \triangleq \{x : h(x) \geq \gamma\}$ είναι κυρτό.

Απόδειξη: Έστω $x_1, x_2 \in$ . Θα δείξουμε ότι $\theta x_1 + (1-\theta)x_2 \in H_\gamma$ για όλα τα $\theta \in [0,1]$. Από τον ορισμό της κοίλης συνάρτησης, έχουμε ότι

$$h(\theta x_1 + (1-\theta)x_2) \geq \theta h(x_1) + (1-\theta)h(x_2) \geq \theta\gamma + (1-\theta)\gamma = \gamma$$

επομένως $\theta x_1 + (1-\theta)x_2 \in H_\gamma$ □

Ανάλογη πρόταση ισχύει για τα σύνολα $\{x : h(x) \leq \gamma\}$ όταν η $h$ είναι κυρτή.

**Θεώρημα 4.5**

Έστω $\phi_i(x) \geq 0$, $i = 1, 2, ..., m$ και οι $\phi_i$ είναι κοίλες συναρτήσεις-περιορισμοί ενός προβλήματος μη-γραμμικού προγραμματισμού. Τότε η εφικτή περιοχή

$$C = \{x : \phi_i(x) \geq 0,\ i = 1, ..., m\}$$

είναι κυρτό σύνολο.

Απόδειξη: Για την $i$-οστή συνάρτηση, το σύνολο $H_\gamma^i$ (βλ. προηγούμενο θεώρημα) με $\gamma = 0$ είναι κυρτό. Επίσης $C = \cap_1^m H_\gamma^i$, οπότε $C$ κυρτό. □

Ανάλογη πρόταση ισχύει για περιορισμούς $\phi_i(x) \leq 0$ όταν οι $\phi_i$ είναι κυρτές.
Συνέπειες των παραπάνω είναι τα εξής:

- Ο ημιχώρος $\{x : a^T x \geq b\}$ είναι κυρτός για όλα τα διανύσματα $a \neq 0$.
- Έστω $A \in \mathbb{R}^{m\times n}$ και $b \in \mathbb{R}^m$. Το σύνολο $\{x : Ax = b\}$ είναι κυρτό.

### 3.4.2 Αναγκαίες συνθήκες για βέλτιστο - Θεώρημα KKT (Karush-Kuhn-Tucker)

Η τυπική μορφή προβλήματος μη-γραμμικού προγραμματισμού με ανισοτικούς περιορισμούς είναι η εξής:

**Πρόβλημα 1** *Έστω $f : \mathbb{R}^n \to \mathbb{R}$ παραγωγίσιμη και κοίλη. $Max.\ f(x)$υ.π. $\phi_i(x) \geq 0$, $i = 1, 2, ..., m$, όπου $\phi_i : \mathbb{R}^n \to \mathbb{R}$ παραγωγίσιμες και κοίλες.*

Σε κάθε σημείο $x$ της εφικτής περιοχής για το παραπάνω πρόβλημα, οι περιορισμοί χωρίζονται σε 2 κατηγορίες:

- Ενεργοί στο $x$: αυτοί για τους οποίους $\phi_i(x) = 0$.
- Μη-ενεργοί στο $x$: αυτοί για τους οποίους $\phi_i(x) > 0$.

Εάν για παράδειγμα το σημείο $x^*$ το οποίο είναι βέλτιστο για το Πρόβλημα 1 είναι εσωτερικό σημείο της εφικτής περιοχής, τότε $\phi_i(x) < 0$ για όλα τα $i = 1, 2, ..., m$, οπότε μπορούμε να αγνοήσουμε όλους τους περιορισμούς (να λύσουμε το πρόβλημα χωρίς αυτούς και να καταλήξουμε στην ίδια βέλτιστη λύση, $x^*$). Εάν γνωρίζαμε εκ των προτέρων ποιοι είναι οι ενεργοί περιορισμοί στο βέλτιστο $x^*$, τότε θα αγνοούσαμε τους υπόλοιπους, και θα λύναμε το πρόβλημα με περιορισμούς ισότητας, χρησιμοποιώντας τη μέθοδο των πολλαπλασιαστών Lagrange που αναπτύξαμε προηγουμένως. Συνήθως όμως είμαστε αναγκασμένοι να φθάσουμε στη λύση προοδευτικά, ενεργοποιώντας μερικούς από τους περιορισμούς κάθε φορά.

Το παρακάτω θεώρημα (των Karush-Kuhn-Tucker) μας δίνει τις αναγκαίες συνθήκες για βέλτιστο της $f$ όταν υπάρχουν περιορισμοί ανισότητας:

**Θεώρημα 4.6**

(KKT) Έστω $f : \mathbb{R}^n \to \mathbb{R}$ παραγωγίσιμη και κοίλη. Για το πρόβλημα $Max.\ f(x)$ υ.π. $\phi_i(x) \geq 0$, $i = 1, 2, ..., m$, όπου $\phi_i : \mathbb{R}^n \to \mathbb{R}$ παραγωγίσιμες και κοίλες: Η Λαγκραντζιανή είναι

$$L = f(x) + \lambda_1\phi_1(x) + \lambda_2\phi_2(x) + ... + \lambda_m\phi_m(x)$$

Οι αναγκαίες συνθήκες για βέλτιστο είναι

$$\frac{\partial L}{\partial x} = 0 \tag{3.30}$$

$$\lambda_i\phi_x(x) = 0, i = 1, ..., m \tag{3.31}$$

$$\phi_i(x) \geq 0 \tag{3.32}$$

$$\lambda_i \geq 0 \tag{3.33}$$

Για να αποδείξουμε το θεώρημα KKT, χρειαζόμαστε το επόμενο Λήμμα:

**Λήμμα 4.1**

(Farkas) Το διάνυσμα $b$ ικανοποιεί την σχέση $b^T y \geq 0\ \forall\ y : Ay \geq 0$ αν και μόνο αν $\exists\ w \geq 0 : A^T w = b$

Απόδειξη:

Έστω $\exists w \geq 0 : A^T w = b$. Τότε $b^T y = (w^T A)y = w^T Ay$. Αν $Ay \geq 0$, τότε $b^T y \geq 0$ αφού $w \geq 0$.

Για να αποδείξουμε την αντίστροφη πρόταση, έστω $b^T y \geq 0\ \forall\ y : Ay \geq 0$. Σχηματίζουμε το εξής πρόβλημα γραμμικού προγραμματισμού:

$$Min\ Z = b^T y,\ \text{υ.π. } Ay \geq 0$$

με τις μεταβλητές του διανύσματος $y$ ελεύθερες. Το πρόβλημα αυτό έχει ως δυϊκό του το

$$Max\ Z = w^T 0 = 0,\ \text{υ.π.}\ A^T w = b,\ w \geq 0$$

Παρατηρούμε ότι το αρχικό ΠΓΠ έχει εφικτή λύση την $y = 0$. Από την θεωρία του γραμμικού προγραμματισμού, είναι γνωστό ότι το δυϊκό πρόβλημα έχει εφικτή λύση αν και μόνο αν έχει εφικτή λύση το αρχικό. Συμπεραίνουμε λοιπόν, ότι υπάρχει $w \geq 0$ τέτοιο ώστε $A^T w = b$. □

**Απόδειξη του θεωρήματος ΚΚΤ**:
Η (3.32) είναι αυτονόητη, καθώς πρέπει να ισχύουν οι περιορισμοί που δόθηκαν με το πρόβλημα. Για την (3.30), έστω οτι ζητάμε να μεγιστοποιήσουμε την $f$. Θα πρέπει να δείξουμε ότι η (3.30) ισχύει στο σημείο $x^*$ αν και μόνο αν για κάθε "αρκετά μικρή" $\epsilon$-γειτονιά του $x^*$, οι τιμές της $f$ σε κάθε σημείο της δεν είναι μεγαλύτερες από $f(x^*)$:

$$\exists \epsilon_0 : \forall\ \epsilon < \epsilon_0, \|x - x^*\| \leq \epsilon \Rightarrow f(x) \leq f(x^*)$$

Θέτουμε λοιπόν $d \in \mathbb{R}^n$, $||d|| \leq \epsilon$ ένα "μικρό" διάνυσμα, προκειμένου να εξετάσουμε τις τιμές της $f$ στο $x^* + d$. Εκ πρώτης όψεως φαίνεται πως πρέπει να δείξουμε οτι για $\epsilon > 0$ αρκετά μικρό, $f(x^* + d) \leq f(x^*)$, οπότε με βάση το συλλογισμό που χρησιμοποιήσαμε προηγουμένως για τη μεγιστοποίηση συναρτήσεων (ανάπτυξη της $f(x^* + d)$ σε σειρά Taylor γύρω από το $x^*$) θα λέγαμε ότι πρέπει να ισχύει

$$\left( \frac{\partial f}{\partial x}\bigg|_{x^*} \right)^T d \leq 0 \tag{3.34}$$

(δηλαδή η μεταβολή της $f$ σε πρώτη τάξη λόγω της "μετατόπισης" του $x^*$ προς την κατεύθυνση του $d$ να μην βελτιώνει την τιμή της συνάρτησης).

Στην πραγματικότητα όμως, αρκεί να ισχύει κάτι λιγότερο αυστηρό: η (3.34) αρκεί να ισχύει *μόνο* για τις μετατοπίσεις $d$ για τις οποίες

$$\left( \frac{\partial \phi_i}{\partial x}\bigg|_{x^*} \right)^T d \geq 0, \quad i \in \{1, ..., m\},$$

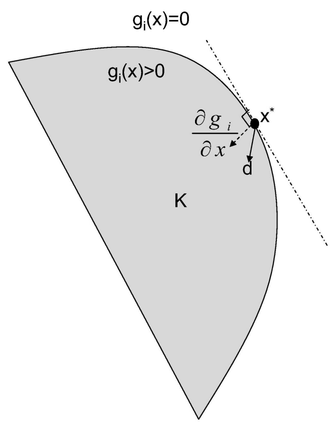
δηλαδή για τα $d$ για τα οποία ο πρώτης τάξης όρος στη σειρά Taylor του κάθε περιορισμού δεν προκαλεί μεταβολή στην $\phi(x^*) \to \phi(x^* + d)$ τέτοια ώστε να παραβιάζεται ο περιορισμός στο $x^* + d$.

Η τελευταία σχέση ζητά το $x^*$ να μετατοπίζεται μόνο προς κατευθύνσεις $d$ προς τις οποίες οι περιορισμοί να εξακολουθούν να ισχύουν (Σχ. 3.3). Για παράδειγμα, σε περίπτωση που στο $x^*$ είναι ενεργοί όλοι οι περιορισμοί, δεν είναι δυνατό να έχουμε

$$\left( \frac{\partial \phi_i}{\partial x}\bigg|_{x^*} \right)^T d < 0$$

για κάποιο $i$, διότι τότε για αρκετά μικρό $||d||$, το σημείο $x^* + d$ θα ήταν έξω από την εφικτή περιοχή.

Στο λήμμα του Farkas, αντικαθιστούμε για το διάνυσμα $w$ το $\lambda$, για το $y$ το $d$, για το $b$ το $-\frac{\partial f}{\partial x}|_{x^*}$, και για τον πίνακα $A^T$ τον $\left[ \frac{\partial \phi_1}{\partial x}|_{x^*}\ \frac{\partial \phi_2}{\partial x}|_{x^*}\ \cdots\ \frac{\partial \phi_m}{\partial x}|_{x^*} \right]$. Άρα, έχουμε άμεσα ότι




**Σχήμα 3.3.** Αν ο περιορισμός $\phi_i$ είναι ενεργός (το βέλτιστο επιτυγχάνεται όταν $\phi_i(x) = 0$) τότε οι μόνες δυνατές μετατοπίσεις $d$ είναι αυτές που δεν θέτουν το $x + d$ εκτός εφικτής περιοχής, δηλαδή τα διανύσματα $d$ τα οποία έχουν μη-αρνητικό εσωτερικό γινόμενο με το διάνυσμα κλίσης της $\phi_i(x)$

$$\left( \left. \frac{\partial f}{\partial x} \right|_{x^*} \right)^T d \leq 0 \; \forall \, d : \left[ \; \tfrac{\partial \phi_1}{\partial x}|_{x^*} \; \tfrac{\partial \phi_2}{\partial x}|_{x^*} \; \cdots \; \tfrac{\partial \phi_m}{\partial x}|_{x^*} \; \right]^T d \geq 0 \tag{3.35}$$

αν και μόνο αν

$$\exists \lambda \geq 0 : \left[ \frac{\partial \phi_1}{\partial x} \frac{\partial \phi_2}{\partial x} \cdots \frac{\partial \phi_m}{\partial x} \right] \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} = - \left. \frac{\partial f}{\partial x} \right|_{x^*}$$

Η τελευταία σχέση (παρατηρώντας ότι η αριστερή πλευρά της είναι γραμμικός συνδυασμό των στηλών $\partial \phi_i / \partial x$ με συντελεστές τα $\lambda_i$) μπορεί να γραφεί ως εξής:

$$\exists \lambda \geq 0 : \left. \frac{\partial f}{\partial x} \right|_{x^*} + \sum_1^m \lambda_i \left. \frac{\partial \phi_i}{\partial x} \right|_{x^*} = 0$$

το οποίο ισοδυναμεί με τις (3.30) και (3.33), ή

$$\exists \lambda \geq 0 : \frac{\partial L}{\partial x} = 0$$

Τέλος, για την 3.31, σημειώνουμε ότι στο τοπικό βέλτιστο θα έχουμε είτε $\phi_i(x) = 0$ (η βέλτιστη λύση “προσκρούει” σε κάποιον από τους περιορισμούς), ή $\lambda_i = 0$ (διότι ο περιορισμός είναι μη-ενεργός και δεν χρειάζεται να τον λάβουμε υπόψη προσθέτοντάς στον στην Λαγκρανζιανή). Σε κάθε περίπτωση, $\lambda_i \phi_i = 0$, οπως δηλαδή διατυπώθηκε και στην πρόταση του θεωρήματος ΚΚΤ. $\square$

Οι συνθήκες του θεωρήματος KKT είναι και ικανές:

**Θεώρημα 4.7**

Για το Πρόβλημα 1, έστω η ότι συνάρτηση $f : C \to \mathbb{R}$ κοίλη, το $C$ κυρτό, και οι $f$, $\phi_i$, $i = 1, ..., m$ παραγωγίσιμες, ενώ οι $\phi_i$ είναι κοίλες. Αν ένα σημείο $x^*$ ικανοποιεί τις συνθήκες KKT, τότε το $x^*$ είναι βέλτιστο.

Απόδειξη: Έστω οποιοδήποτε $y$ εφικτό. Θα δείξουμε ότι $f(y) \leq f(x^*)$. Εφ'όσον οι $\phi_i$ είναι κοίλες, η εφικτή περιοχή είναι κυρτή (όπως δείξαμε προηγουμένως). Θέτουμε $d = y - x^*$. Για τους ενεργούς περιορισμούς, θα έχουμε

$$\left(\frac{\partial \phi_i}{\partial x}\right)^T d \geq 0$$

διότι στην αντίθετη περίπτωση, για $||d||$ αρκετά μικρό, $\phi_i(y) = \phi_i(x^*+d) \approx \phi(x^*)+\left(\frac{\partial \phi_i}{\partial x}\right)^T d < 0$ (βλ. σειρά Taylor), και άρα το $y$ θα βρίσκεται έξω από την εφικτή περιοχή. Από το θεώρημα KKT, έχουμε ότι

$$\left(\left.\frac{\partial f}{\partial x}\right|_{x^*}\right)^T d \leq 0 \tag{3.36}$$

Επίσης, επειδή η $f$ είναι κοίλη, ισχύει (Θεώρημα 4.2)

$$f(y) - f(x^*) \leq \left(\left.\frac{\partial f}{\partial x}\right|_{x^*}\right)^T (y - x^*) = \left(\left.\frac{\partial f}{\partial x}\right|_{x^*}\right)^T d \leq 0.$$

Άρα το $x^*$ είναι βέλτιστο. $\square$

## 3.5 Παραδείγματα

Ακολουθεί μια σειρά από παραδείγματα με αυξανόμενο βαθμό δυσκολίας.

### Παράδειγμα 6

Έστω ότι θέλουμε να μεγιστοποιήσουμε την

$$f(x) = -(x_1 - 1/2)^2 - 2(x_2 - 1/2)^2,$$

υ.π. $2x_1 + x_2 \leq 2$, $x_1, x_2 \geq 0$. Ξεκινάμε μετασχηματίζοντας τους περιορισμούς στη μορφή που απαιτεί το θεώρημα KKT:

$$\phi_1(x) = -2x_1 - x_2 + 2 \geq 0$$

$$\phi_2(x) = x_1 \geq 0$$

$$\phi_3(x) = x_2 \geq 0$$

Στη συνέχεια ελέγχουμε ότι η εφικτή περιοχή του προβλήματος είναι κυρτή (αντίστοιχα, ότι όλοι οι $\phi_i(x) \geq 0$ είναι κοίλες συναρτήσεις. Όντως, οι $\phi_1, \phi_2, \phi_3$ είναι οπωσδήποτε κοίλες, ως γραμμικές. Τέλος, ελέγχουμε το κατά πόσο η αντικειμενική συνάρτηση $f$ είναι κοίλη η κυρτή υπολογίζοντας τον Εσσιανό της πίνακα:

$$\frac{\partial^2 f}{\partial x \partial x} = \begin{bmatrix} -2 & 0 \\ 0 & -4 \end{bmatrix}.$$

Ο παραπάνω πίνακας είναι προφανώς αρνητικά ορισμένος, άρα η $f$ είναι κοίλη. Συνεπώς, το παραπάνω πρόβλημα μπορεί να λυθεί με χρήση του θεωρήματος ΚΚΤ.

Σχηματίζουμε τη Λαγκραντζιανή

$$L = -(x_1 - 1/2)^2 - 2(x_2 - 1/2)^2 + \lambda_1(-2x_1 - x_2 + 2) + \lambda_2(x_1) + \lambda_3(x_2)$$

και θα πρέπει να λύσουμε το σύστημα που προκύπτει από τις συνθήκες ΚΚΤ:

$$\begin{aligned} \partial L/\partial x &= 0 && (3.37) \\ \lambda_1(-2x_1 - x_2 + 2) &= 0 \\ \lambda_2(x_1) &= 0 \\ \lambda_3(x_2) &= 0 \\ \phi_{1,2,3}(x) &\geq 0 && (3.38) \\ \lambda_{1,2,3} &\geq 0 \end{aligned}$$

Αντί να λύσουμε ταυτόχρονα τις συνθήκες ΚΚΤ, μπορούμε να δοκιμάσουμε την εξής διαδικασία:

- υποθέτουμε αρχικά ότι όλοι οι περιορισμοί είναι ανενεργοί - δηλ. $\lambda_i = 0$ για όλα τα $i$. Αυτό σημαίνει ότι "στοιχηματίζουμε" ότι το βέλτιστο σημείο θα βρίσκεται στο εσωτερικό της εφικτής περιοχής και άρα κανένας περιορισμός δε θα συμμετέχει ουσιαστικά στον προσδιορισμό του βέλτιστου (και γι'αυτό, θέτοντας $\lambda_i = 0$ "αφαιρούμε" τους περιορισμούς από το πρόβλημα). Αν συμβεί οι συνθήκες ΚΚΤ να έχουν εφικτή λύση με $\lambda_i = 0$ αυτό θα σημαίνει και ότι όντως το βέλτιστο βρέθηκε σε εσωτερικό σημείο.
- Αν όμως το σύστημα ΚΚΤ δεν έχει εφικτή λύση με $\lambda_i = 0$ αυτό θα σημαίνει ότι η λύση που βρήκαμε παραβίαζε τουλάχιστον έναν περιορισμό. Σε τέτοια περίπτωση μπορούμε να ενεργοποιήσουμε τον/τους περιορισμό/ούς που παραβιάστηκαν (θέτουμε τις αντίστοιχες $\phi_i(x) = 0$ και τους αντίστοιχους πολλαπλασιαστές $\lambda_i > 0$) και να ξαναλύσουμε το σύστημα ΚΚΤ, επαναλαμβάνοντας τη διαδικασία αυτή έως ότου βρούμε εφικτή λύση.

Αν εφαρμόσουμε την παραπάνω μέθοδο στο πρόβλημά μας, θα θέσουμε αρχικά $\lambda_{1,2,3} = 0$ και

$$L = -(x_1 - 1/2)^2 - 2(x_2 - 1/2)^2$$

$$\partial L/\partial x = \begin{bmatrix} -2(x_1 - 1/2) \\ -4(x_2 - 1/2) \end{bmatrix} = 0$$

το οποίο δίνει άμεσα $x^* = x = [1/2, 1/2]^T$. Παρατηρούμε ότι αυτή η λύση ικανοποιεί όλες τις συνθήκες ΚΚΤ, και άρα η $f$ μεγιστοποιείται στο παραπάνω σημείο $x^*$. Σε αυτή την περίπτωση λοιπόν, το μέγιστο βρέθηκε σε εσωτερικό σημείο της εφικτής περιοχής, και το πρόβλημα (όπως δείχνει ο μηδενισμός των πολλαπλασιαστών $\lambda_i$ μπορεί να λυθεί αγνοώντας τους περιορισμούς.

### Παράδειγμα 7

Έστω ότι αλλάζουμε λίγο την αντικειμενική συνάρτηση του προηγούμενου παραδείγματος σε

$$Max f(x) = -(x_1 - 1)^2 - 2(x_2 - 1)^2$$

, χωρίς να μεταβάλουμε τις συναρτήσεις των περιορισμών:

$$\phi_1(x) = -2x_1 - x_2 + 2 \geq 0, \quad \phi_2(x) = x_1 \geq 0, \quad \phi_3(x) = x_2 \geq 0$$

οπότε η Λαγκραντζιανή μας τώρα θα είναι

$$L = -(x_1 - 1)^2 - 2(x_2 - 1)^2 + \lambda_1(-2x_1 - x_2 + 2) + \lambda_2(x_1) + \lambda_3(x_2)\ddot{}.$$

Επαναλαμβάνοντας την διαδικασία που προτάθηκε στο προηγούμενο παράδειγμα, δοκιμάζουμε αρχικά το κατά πόσο μπορεί να βρεθεί το μέγιστο σε εσωτερικό σημείο της εφικτής περιοχής, δηλαδή θέτουμε $\lambda_1 = \lambda_2 = \lambda_3 = 0$, και λύνουμε τις συνθήκες ΚΚΤ:

$$L = -(x_1 - 1)^2 - 2(x_2 - 1)^2$$

$$\begin{aligned} \partial L/\partial x = \begin{bmatrix} -2(x_1 - 1) \\ -4(x_2 - 1) \end{bmatrix} &= 0 \qquad (3.39) \\ 0 \cdot (-2x_1 - x_2 + 2) &= 0 \\ 0 \cdot (x_1) &= 0 \\ 0 \cdot (x_2) &= 0 \\ \phi_{1,2,3}(x) &> 0 \qquad (3.40) \end{aligned}$$

Αυτή τη φορά, οι πρώτες δύο εξισώσεις δίνουν $x_1 = 1, x_2 = 1$, το οποίο όμως δεν είναι εφικτό καθώς, όπως εύκολα μπορούμε να διαπιστώσουμε, δεν ικανοποιεί τον περιορισμό $\phi_1(x) \geq 0$.

Για να συνεχίσουμε ενεργοποιούμε τον $\phi_1$ (δηλαδή τον περιορισμό τον οποίο παραβίασε η παραπάνω λύση) και υποθέτουμε ότι $\phi_1(x) = 0$ (άρα $\lambda_1 > 0$), ενώ οι $\phi_2$ και $\phi_3$ παραμένουν ανενεργοί (δηλ. $\lambda_2 = \lambda_3 = 0$). Στη συνέχεια δοκιμάζουμε να λύσουμε εκ νέου τις συνθήκες ΚΚΤ:

$$L = -(x_1 - 1)^2 - 2(x_2 - 1)^2 + \lambda_1(-2x_1 - x_2 + 2)$$

$$\partial L/\partial x = \begin{bmatrix} -2(x_1-1)-2\lambda_1 \\ -4(x_2-1)-\lambda_1 \end{bmatrix} = 0 \tag{3.41}$$

$$\begin{aligned} \lambda_1(-2x_1-x_2+2) &= 0 \\ 0\cdot(x_1) &= 0 \\ 0\cdot(x_2) &= 0 \\ \phi_1(x) = -2x_1-x_2+2 &= 0 \\ \phi_2(x) = x_1 &> 0 \\ \phi_3(x) = x_2 &> 0 \end{aligned}$$

Αυτή τη φορά, το (γραμμικό) σύστημα που προκύπτει από τις δύο πρώτες εξισώσεις μαζί και με τον ενεργό περιορισμό, είναι

$$\begin{bmatrix} -2(x_1-1)-2\lambda_1 \\ -4(x_2-1)-\lambda_1 \end{bmatrix} = 0 \tag{3.42}$$

$$-2x_1-x_2+2=0 \tag{3.43}$$

και λύνεται εύκολα (π.χ. με αντικατάσταση) για να δώσει $x_1 = 5/9$, $x_2 = 8/9$, $\lambda_1 = 4/9$. Το σημείο αυτό είναι προφανώς εφικτό (δίνει $\phi_1 = 0$, $\phi_{2,3} > 0$, και $\lambda_i \geq 0$) άρα, βασιζόμενοι στο θεώρημα KKT συμπεραίνουμε ότι αποτελεί το μέγιστο της $f$ υπό τους δεδομένους περιορισμούς.

### Παράδειγμα 8

Έστω $x = [x_1, x_2]^T$. Min $\psi(x) = -3x_1 + x_2^2/2$, υ.π.

$$\begin{aligned} &x_1^2+x_2^2-1 \leq 0 \\ &x_1 \geq 0 \\ &x_2 \geq 0 \end{aligned}$$

Σύμφωνα με τους περιορισμούς, το $x$ θα βρίσκεται σε μια περιοχή που ορίζεται από την τομή του άνω-δεξιά τεταρτημορίου και ενός δίσκου με ακτίνα 1.

Αρχικά, μετασχηματίζουμε το πρόβλημα στην τυπική μορφή, ώστε να εφαρμόσουμε το θεώρημα KKT:
Max $-\psi(x) = 3x_1 - x_2^2/2$, υ.π.

$$\begin{aligned} \phi_1(x) &= -x_1^2-x_2^2+1 \geq 0 \\ \phi_2(x) &= x_1 \geq 0 \\ \phi_3(x) &= x_2 \geq 0 \end{aligned}$$

Επιβεβαιώστε (από τον πίνακα των δευτερων παραγώγων των $\phi_i(x)$)) ότι α) οι $\phi_{1,2,3}$ είναι κοίλες, άρα η εφικτή περιοχή είναι κυρτό σύνολο, και (απο τον πίνακα των δευτερων παραγώγων της $\psi(x)$) ότι β) η $\psi(x)$ είναι κυρτή στην εφικτή περιοχή. Άρα η $-\psi$ είναι κοίλη συνάρτηση της οποίας ζητάμε το μέγιστο σε κυρτή εφικτή περιοχή, και το παραπάνω πρόβλημα εμπιπτει στο θεώρημα KKT.

Σχηματίζουμε τη Λαγκραντζιανή

$$L = (3x_1 - x_2^2/2) + \lambda_1(-x_1^2 - x_2^2 + 1) + \lambda_2(x_1) + \lambda_3(x_2)$$

και γράφουμε τις συνθήκες KKT:

$$\begin{aligned} \partial L/\partial x = \begin{bmatrix} 3 - 2\lambda_1 x_1 + \lambda_2 \\ -x_2 - 2\lambda_1 x_2 + \lambda_3 \end{bmatrix} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \lambda_1(-x_1^2 - x_2^2 + 1) &= 0 \\ \lambda_2(x_1) &= 0 \\ \lambda_3(x_2) &= 0 \\ \lambda_{1,2,3} &> 0 \end{aligned}$$

Αυτή τη φορά, ας προσπαθήσουμε να λύσουμε το παραπάνω σύστημα ως προς $x_1, x_2, \lambda_{1,2,3}$. Παρατηρούμε ότι αν $-x_1^2 - x_2^2 + 1 \neq 0$, τότε $\lambda_1 = 0$ οπότε $\lambda_2 = -3 < 0$ το οποίο δεν είναι δυνατό. Άρα $\lambda_1 \neq 0$ and $\phi_1 = 0$ (ενεργός).

Επίσης, αν $\lambda_2 \neq 0$ τότε $x_1 = 0$ και $3 - 2\lambda_1 0 + \lambda_2 = 0 \Rightarrow \lambda_2 < 0$, το οποίο δεν ικανοποιεί τις συνθήκες KKT. Άρα, $\lambda_2 = 0$.

Αν $x_2 \neq 0$ τότε $\lambda_3 = 0$, $x_2(-1 - 2\lambda_1) = 0 \Rightarrow x_2 = 0$. Όμως, υποθέσαμε ότι $x_2 \neq 0$, άρα πρέπει να ισχύει $x_2 = 0$.

Τελικά από τις $\phi_1 = 0$ και $x_2 = 0$ έχουμε $x_1 = \pm 1$. Επειδή $\phi_2(x) = x_1 \geq 0$, επιλέγουμε $x_1 = 1$.

Βρήκαμε λοιπόν ότι $x = [1, 0]^T$ και $\lambda = [3/2, 0, 0]^T$ ικανοποιούν τις συνθήκες του θεωρήματος KKT, άρα έχουμε βρεί μέγιστο για την $-\psi(x)$ ή, αντίστοιχα, ελάχιστο για την αρχική $\psi(x)$, όπως ζητήθηκε (στην περίπτωση αυτή τυγχάνει να είναι και ολικό ελάχιστο).

### Παράδειγμα 9

Να ελαχιστοποιήσετε και να μεγιστοποιήσετε την $f = x^3 - x + y^2$, υ.π. $4x^2 + y^2 \leq 4$, $x, y \geq 0$.

Αφού μετατρέψουμε τους περιορισμούς στη μορφή

$$\begin{aligned} \phi_1(x, y) &= -4x^2 - y^2 + 4 \geq 0 \\ \phi_2(x, y) &= x \geq 0 \\ \phi_3(x, y) &= y \geq 0 \end{aligned}$$

ελέγχουμε αν η $f$ είναι κοίλη ή κυρτή. Στην προκειμένη περίπτωση είναι, γιατί ο πίνακας των δεύτερων μερικών παραγώγων της $f$ είναι θετικά ημι-ορισμένος, ενώ η $f$ είναι τρίτου βαθμού. Επίσης, οι περιορισμοί είναι κοίλοι (υπολογίστε τον πίνακα των δεύτερων παραγώγων της $\phi_1$ και ελέγξτε ότι είναι αρνητικά ορισμένος, οι $\phi_2$, $\phi_3$ είναι κοίλοι ως γραμμικοί), άρα η εφικτή περιοχή είναι κυρτή.

**Ελαχιστοποίηση της** $f$: Εφ'όσον η $f$ είναι κυρτή, το πρόβλημα $Minf$ είναι ισοδύναμο με το $Max(-f)$ με την $-f$ να είναι κοίλη. Μπορούμε λοιπόν να χρησιμοποιήσουμε το θεώρημα ΚΚΤ μεγιστοποιώντας την $g(x,y) = -f(x,y)$.

Ξεκινάμε δοκιμάζοντας να μεγιστοποιήσουμε την $g(x,y)$ χωρίς περιορισμούς, σε περίπτωση που το ελάχιστο συμβεί να βρίσκεται σε εσωτερικό σημείο της εφικτής περιοχής (όλοι οι περιορισμοί μη-ενεργοί). Θέτουμε:

$$\partial g/\partial x = \partial g/\partial y = 0$$

οι οποίες μας δίνουν

$$-3x^2 + 1 = 0, \quad -2y = 0$$

ή $x = 1/\sqrt{3}, y = 0$, και $f(1/\sqrt{3}, 0) = -2\sqrt{3}/9$. Το σημείο αυτό ικανοποιεί τους περιορισμούς και όλες τις υπόλοιπες συνθήκες ΚΚΤ άρα είναι μέγιστο της $g(x) = -f(x)$ και συνεπώς ελάχιστο της $f(x)$.

**Μεγιστοποίηση της** $f$: Εφ'όσον και η $f$ και η εφικτή περιοχή είναι κυρτές, το μέγιστο θα βρίσκεται σε ακρότατο της εφικτής περιοχής. Τα ακρότατα είναι:

- Το σημείο $(0,0)$ όπου έχουμε $f = 0$.
- Τα σημεία της καμπύλης που ορίζεται από τον περιορισμό $4x^2+y^2 = 4$, για $x, y \geq 0$. Πρέπει λοιπόν να μεγιστοποιήσουμε την $f$ υ.π. $\phi_1(x,y) = 4x^2 + y^2 - 4 = 0$, και να συγκρίνουμε το μέγιστο με τις τιμές στα υπόλοιπα ακρότατα.

Για να προχωρήσουμε, σχηματίζουμε την Λαγκραντζιανή

$$L = x^3 - x + y^2 + \lambda(4 - 4x^2 - y^2)$$

Οι αναγκαίες συνθήκες για μέγιστο είναι

$$\partial L/\partial x = \partial L/\partial y = 0$$

ή

$$3x^2 - 1 - 8\lambda x = 0 \tag{3.44}$$

$$2y - 2y\lambda = 0 \tag{3.45}$$

Οι τελευταίες 2 εξισώσεις, μαζί με τον περιορισμό $4x^2 + y^2 = 4$ δίνουν ένα σύστημα τριών εξισώσεων με τρεις αγνώστους. Απο την $\partial L/\partial y = 0$, βλέπουμε ότι $\lambda = 1$ ή $y = 0$.

Αν υποθέσουμε ότι $\lambda = 1$, και λύσουμε για $x$ από τη σχέση $3x^2 - 1 - 8x = 0$, θα δούμε ότι καμία από τις λύσεις που προκύπτουν για το $x$ δεν είναι εφικτή (θα παραβιάζεται ο περιορισμός $\phi_1$ ή η μη-αρνητικότητα του $x$). Άρα οπωσδήποτε $y = 0$, το οποίο μας δίνει (από την εξίσωση του περιορισμού $\phi_1 = 0$) ότι $x = 1$, και (από την $\partial L/\partial x = 0$), $\lambda = 1/4$.

Το σημείο που βρήκαμε είναι στάσιμο για την $L$ αλλά θα πρέπει να διαπιστώσουμε αν αντιστοιχεί στο ζητούμενο μέγιστο ή αν είναι τοπικό ελάχιστο. Για το σκοπό αυτό, σχηματίζουμε τον Εσσιανό πίνακα της $L$ στο σημείο $(1,0),\ \lambda = 1/4$:

$$H = \frac{\partial^2 f}{\partial x \partial x} = \begin{bmatrix} 6x - 8\lambda & 0 \\ 0 & 2 - 2\lambda \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 3/2 \end{bmatrix} > 0$$

και στη συνέχεια τον επαυξημένο Εσσιανό

$$\bar{H} = \begin{bmatrix} 0 & 8x & 2y \\ 8x & 6x-8\lambda & 0 \\ 2y & 0 & 2-2\lambda \end{bmatrix} = \begin{bmatrix} 0 & 8 & 0 \\ 8 & 4 & 0 \\ 0 & 0 & 3/2 \end{bmatrix}.$$

Το πρόβλημά μας έχει $n = 2$ μεταβλητές και $m = 1$ περιορισμό, άρα μας ενδιαφέρει το πρόσημο της $n - m = 1$ μεγαλύτερης ορίζουσας του $\bar{H}$. Θα είναι $|\bar{H}| = -96 < 0$ το οποίο μας οδηγεί στο συμπέρασμα ότι το σημείο που βρήκαμε είναι τοπικό *ελάχιστο* και όχι μέγιστο όπως ζητήθηκε.

Αυτό που συμβαίνει είναι ότι το μέγιστο επί της καμπύλης $4x^2 + y^2 = 4$ βρίσκεται στο $(x, y) = (-1, 0)$, σημείο το οποίο αποκλείσαμε κατά την επίλυση του τελευταίου συστήματος καθ'ότι αντιστοιχεί σε σημείο εκτός εφικτής περιοχής. Φαίνεται λοιπόν, ότι η $f$ αυξάνει κατά μήκος της καμπύλης που ορίζεται από τον περιορισμό $4x^2 + y^2 = 4$, αν φανταστούμε ότι ξεκινάμε από το σημείο $(1, 0)$ και κινούμαστε κατά μήκος της έλλειψης $\phi_1 = 0$. Με δεδομένο ότι η $f$ δεν έχει τοπικό ελάχιστο μεταξύ του σημείου $(1, 0)$ και του $(2, 0)$ τα οποία και οριοθετούν το εφικτό τμήμα της καμπύλης, το μέγιστο θα βρίσκεται στα ακρότατα σημεία αυτού του τμήματος, το οποίο μπορούμε να θεωρήσουμε ως μια "νέα" εφικτή περιοχή $\{(x, y) : 4-4x^2-y^2 = 0, x, y \geq 0\}$. Τα ακρότατα είναι ακριβώς τα παραπάνω σημεία δηλαδή $(0, 2)$ με $f(0, 2) = 4$, και $(1, 0)$ με $f(1, 0) = 0$, και τελικά το μέγιστο βρίσκεται στο σημείο $(0, 2)$.

## 3.6 Υπολογιστική Βελτιστοποίηση

Είναι φανερό από τα παραπάνω, ότι η βελτιστοποίηση συνάρτησης (με ή χωρίς περιορισμούς) τελικά μας οδηγεί στην εύρεση ριζών κάποιας εξίσωσης η οποία δίνεται από τις αναγκαίες συνθήκες για βέλτιστο (π.χ. $\partial L/\partial x = 0$). Όπως είναι γνωστό, αν η εν λόγω εξίσωση είναι πέμπτου ή μεγαλύτερου βαθμού, δεν υπάρχει λύση σε κλειστή μορφή, και αναγκαστικά καταφεύγουμε σε υπολογιστικές μεθόδους.

### 3.6.1 Διχοτόμηση

Η απλούστερη ίσως υπολογιστική μέθοδος για την εύρεση τοπικού μεγίστου συνάρτησης είναι αυτή της *διχοτόμησης*. Δεδομένου ότι η $f : \mathbb{R} \to \mathbb{R}$ είναι κοίλη (ανάλογα αποτελέσματα ισχύουν όταν η $f$ είναι κυρτή), με μέγιστο στο $x^*$, θα έχουμε

$$df/dx > 0 \text{ άν } x < x^*$$

και

$$df/dx < 0 \text{ άν } x > x^*$$

Για να βρούμε το $x^*$ (τις ρίζες τις εξίσωσης $df/df = 0$) ακολουθούμε τον εξής αλγόριθμο:

1. Επιλέγουμε $\epsilon > 0$ το οποίο ορίζει την ακρίβεια με την οποία απαιτούμε να προσδιοριστεί το βέλτιστο.

2. Βρίσκουμε 2 σημεία $\bar{x}, \underline{x}$ τέτοια ώστε

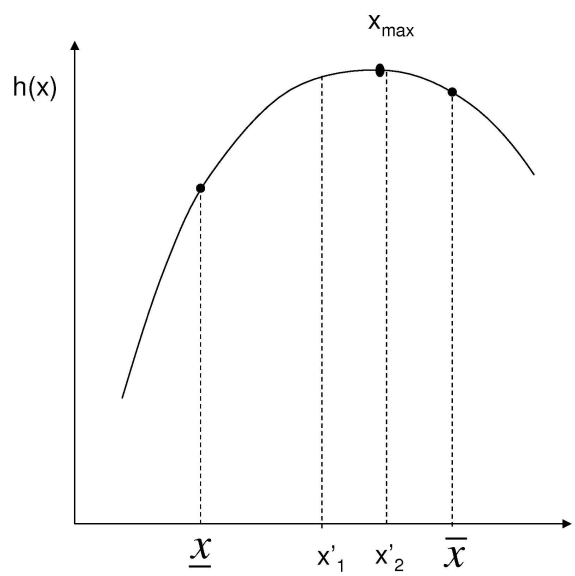
$$\left.\frac{df}{dx}\right|_{\underline{x}} > 0 \tag{3.46}$$

$$\left.\frac{df}{dx}\right|_{\bar{x}} < 0 \tag{3.47}$$

   οπότε $\underline{x} \leq x^* \leq \bar{x}$.
3. Θέτουμε $x' = \frac{\underline{x}+\bar{x}}{2}$.
4. Αν $\bar{x} - \underline{x} \leq 2\epsilon$ τερματίζουμε τον αλγόριθμο.
5. Αν $\left.\frac{df}{dx}\right|_{x'} \geq 0$ τότε θέτουμε $\underline{x} = x'$
6. Αν $\left.\frac{df}{dx}\right|_{x'} \leq 0$ τότε θέτουμε $\bar{x} = x'$
7. Επαναλαμβάνουμε τη διαδικασία από το βήμα 3 του αλγορίθμου.

Όταν ο αλγόριθμος τερματίσει, έχουμε βρει σημείο $x'$ το οποίο απέχει το πολύ $\epsilon$ από το βέλτιστο.

Όπως φαίνεται και από τη μέθοδο με την οποία επιλέγεται το σημείο $x'$ κάθε φορά, ο αλγόριθμος προβαίνει σε διχοτόμηση του διαστήματος $[\underline{x}, \bar{x}]$.




**Σχήμα 3.4.** Εφαρμογή της μεθόδου διχοτόμησης για την εύρεση τ.μεγίστου. Φαίνονται τα αρχικά σημεία $\underline{x}$ και $\bar{x}$, καθώς και οι μέσοι όροι $x'$ που προκύπτουν από τις δύο πρώτες επαναλήψεις του αλγορίθμου.

### 3.6.2 Μέθοδος Newton

Η μέθοδος της διχοτόμησης είναι απλό να εφαρμοστεί αλλά συνήθως συγκλίνει αργά (η κάθε επανάληψη μειώνει τη διαφορά $\bar{x} - \underline{x}$ κατά 1/2 πάντοτε. Αυτό συμβαίνει γιατί οι πληροφορίες που χρησιμοποιεί ο αλγόριθμος είναι περιορισμένες (μόνο η πρώτη παράγωγος της $f$). Η μέθοδος του Newton, την οποία παρουσιάζουμε στη συνέχεια, έρχεται να βελτιώσει την κατάσταση χρησιμοποιόντας και τη δεύτερη παράγωγο της $f$.

Η μέθοδος του Newton βασίζεται στον διαδοχικό υπολογισμό μιας σιράς από σημεία $x_i$, $i = 0, 1, 2, ...$, η οποία συγκλίνει στο $x^*$. Έστω $x_i$ ένα από αυτά τα σημεία. Χρησιμοποιόντας τη σειρά Taylor για την $f$ στο σημείο $x_i$, γράφουμε

$$f(x_{i+1}) \approx f(x_i) + f'(x_i)(x_{i+1} - x_i) + \frac{f''(x_i)}{2}(x_{i+1} - x_i)^2 \tag{3.48}$$

όπου $f'$, $f''$ είναι οι πρώτη και δεύτερη παράγωγος αντίστοιχα, και επιλέξαμε να αγνοήσουμε τους όρους με τάξη μεγαλύτερη του 2. Εφ'όσον θέλουμε να μεγιστοποιήσουμε την $f$, θα υπολογίσουμε το σημείο $x_{i+1}$ το οποίο μεγιστοποιεί την (3.48). Το αποτέλεσμα δεν θα είναι ίσο με το $x^*$ (η (3.48) είναι δευτεροβάθμια προσέγγιση της $f$) αλλά θα βελτιώνει το $x_i$.

Έτσι, θέτουμε την παράγωγο της (3.48) ώς προς $x_{i+1}$ ίση με μηδέν, και λύνουμε για το $x_{i+1}$:

$$0 = f'(x_i) + f''(x_i)(x_{i+1} - x_i) \Rightarrow \tag{3.49}$$

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)} \tag{3.50}$$

Η τελευταία εξίσωση, ξεκινώντας από οποιοδήποτε σημείο $x_0$ ορίζει την σειρά $x_1, x_2, ...$ την οποία υπολογίζουμε έως ότου $|x_{i+1} - x_i| \leq \epsilon$. Όπως αναφέραμε, η μέθοδος του Newton γενικά συγκλίνει γρηγορότερα από την μέθοδο της διχοτόμησης. Ο ρυθμός σύγκλισης εξαρτάται από το βαθμό της $f$ καθώς όσο μεγαλύτερος είναι αυτός τόσο πιο ανακριβής είναι η προσέγγιση της $f$ στην (3.48).

Σε περίπτωση που $f : \mathbb{R}^n \to \mathbb{R}$, η εξίσωση που αντιστοιχεί στην (3.50) είναι

$$x_{i+1} = x_i + \left( \left.\frac{\partial^2 f}{\partial x \partial x}\right|_{x_i} \right)^{-1} \left( \left.\frac{\partial f}{\partial x}\right|_{x_i} \right) \tag{3.51}$$

Αξίζει να σημειώσουμε ότι σε περίπτωση που η $f$ είναι κοίλη, όπως υποθέσαμε παραπάνω, τόσο η διχοτόμηση όσο και η μέθοδος του Newton συγκλίνουν στο *ολικό μέγιστο* της $f$. Αν η $f$ δεν είναι κοίλη, η σύγκλιση γίνεται προς τοπικό μέγιστο.

### 3.6.3 Αλγόριθμοι αναζήτησης με βάση την κλίση της $f$

Όταν $f$ είναι συνάρτηση διανυσματικής μεταβλητής, η αναζήτηση σημείου $x^*$ στο οποίο επιτυγχάνεται τοπικό μέγιστο γίνεται με μεθόδους οι οποίες βασίζονται στην "ακολούθηση της κλίσης" της $f$ από κάποιο αρχικό σημείο, προς άλλα τα στα οποία η $f$ αυξάνεται.

Πιο συγκεκριμένα, αρχίζουμε με οποιοδήποτε σημείο $x_0$ και υπολογίζουμε την κλίση της $f$ στο $x_0$, $\partial f/\partial x$. Η κλίση δίνει τη διεύθυνση προς την οποία πρέπει να μετατοπιστεί το $x_0$ (κατά $t\,\frac{\partial f}{\partial x}\Big|_{x_i}$ όπου $t > 0$) ώστε να αυξηθεί η τιμή της $f$ στο μέγιστο δυνατό βαθμό (δεδομένης της απόστασης μετατόπισης). Είναι φανερό ότι αν η μετατόπιση είναι αρκετά μικρή, το νέο σημείο $x_1 = x_0 + t\,\frac{\partial f}{\partial x}\Big|_{x_i}$ θα έχει $f(x_1) > f(x_0)$. Το ερώτημα είναι ποιο πρέπει να είναι το μέγεθος της μετατόπισης στην κατεύθυνση $\partial f/\partial x$ ώστε να μεγιστοποιείται η αντίστοιχη αύξηση της $f$.

Έχουμε λοιπόν ένα μονοδιάστατο πρόβλημα βελτιστοποίησης: max $f\left(x_i + t\,\frac{\partial f}{\partial x}\Big|_{x_i}\right)$ ως προς $t$, το οποίο μπορεί να λυθεί χρησιμοποιόντας (για παράδειγμα) τη μέθοδο του Newton ή διχοτόμηση. Έστω ότι $t^*$ είναι η τιμή του $t$ η οποία μεγιστοποιεί την $f\left(x_i + t\,\frac{\partial f}{\partial x}\Big|_{x_i}\right)$. Τότε θέτουμε

$$x_{i+1} = x_i + t^*\,\frac{\partial f}{\partial x}\bigg|_{x_i}$$

και επαναλαμβάνουμε την διαδικασία έως ότου $\left|\frac{\partial f}{\partial x}\right| \leq \epsilon$, οπότε και έχουμε βρει σημείο το οποίο απέχει το πολύ $\epsilon$ από το $x^*$.

# Α΄

# Βασικά στοιχεία Γραμμικής Άλγεβρας

## Α΄.1 Σημειογραφία

Γράφουμε $A \in \mathbb{R}^{n\times m}$ για πίνακα $n$(σειρές) επί $m$(στήλες) του οποίου τα στοιχεία είναι πραγματικοί. Ο $A$ μπορεί να ειδωθεί σαν *γραμμικός τελεστής* που "στέλνει" $m$-διάστατα διανύσματα σε $n$-διάστατα διανύσματα:

$$\begin{aligned} A &: \mathbb{R}^m \to \mathbb{R}^n \\ Ax &= y \end{aligned} \tag{Α΄.1}$$

για $x \in \mathbb{R}^m$ και $y \in \mathbb{R}^n$. Για συγκεκριμένα $A$ και $y$, η εξίσωση $Ax = y$ είναι απλά ένας "συμπαγής" τρόπος για να αναπαραστήσουμε ένα σύστημα από $n$ γραμμικές εξισώσεις με $m$ αγνώστους (τα στοιχεία του $x$).

Αν $A \in \mathbb{R}^{n\times m}$, τότε γράφουμε $a_{ij}$ για τα στοιχεία του $A$, για $i = 1, ..., n$, $j = 1, ..., m$. Δηλαδή:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nm} \end{bmatrix} \tag{Α΄.2}$$

$A^T$ δηλώνει τον *ανάστροφο* του πίνακα $A$. $I_n$ δηλώνει τον μοναδιαίο πίνακα $n \times n$. Για παράδειγμα:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{Α΄.3}$$

Συχνά, το σύμβολο $n$ παραλείπεται, και οι διαστάσεις του $I$ συμπεραίνονται από τα συμφραζόμενα ή τους άλλους όρους της έκφρασης στην οποία χρησιμοποιείται ο $I$.

Ο συμβολισμός $Range(A)$ δηλώνει το *χώρο στηλών* του , δηλαδή το χώρο όλων των δυνατών γραμμικών συνδυασμών των στηλών του .

## Α΄.2 Ορίζουσες και Αντιστροφή Πίνακα

Έστω $A \in \mathbb{R}^{n\times n}$. Αν υπάρχει πίνακας $B$ (επίσης $n \times n$) ώστε

$$AB = BA = I_n \tag{Α΄.4}$$

τότε λέμε οτι ο $B$ είναι ο *αντίστροφος* του $A$ και γράφουμε

$$B = A^{-1} \tag{Α΄.5}$$

Ένας (τετράγωνος) πίνακας $A$ έχει αντίστροφο αν και μόνο αν η *ορίζουσά* του, $|A|$, δεν είναι μηδενική (λέμε ότι ο $A$ είναι *μη ιδιάζων*). Υπάρχουν διάφορες μέθοδοι για τον υπολογισμό της ορίζουσας του πίνακα $A$. Ένας αναδρομικός ορισμός του $|A|$ (γνωστός ως μέθοδος του Cramer) είναι:

$$|A| = a_{11}|D_{11}| - a_{12}|D_{12}| + a_{13}|D_{13}| - \ldots + (-1)^n a_{1n}|D_{1n}| \tag{Α΄.6}$$

όπου $|D_{ij}|$ είναι η ορίζουσα του $(n-1)\times(n-1)$ υπο-πίνακα (γνωστός και σαν "ελάσσων") που σχηματίζεται από τον $A$ με διαγραφή της $i$ σειράς και της $j$ στήλης. Στην παραπάνω εξίσωση, επιλέξαμε να "αναπτύξουμε" την ορίζουσα χρησιμοποιόντας τα στοιχεία $a_{1i}$ της πρώτης σειράς του $A$. Βεβαίως, ο κανόνας του Cramer μπορεί να εφαρμοστεί χρησιμοποιόντας τα στοιχεία οποιασδήποτε σειράς ή στήλης του $A$. Μερικές φορές θα γράφουμε $det(A)$ αντί για $|A|$.

Ένας $n\times n$ πίνακας $A$, έχει ακριβώς $n^2$ ελάσσονες $D_{ij}$ διαστάσεων $(n-1)\times(n-1)$ (Γιατί?). Ο τύπος για τον υπολογισμό του αντίστροφου του $A$ είναι:

$$A^{-1} = \frac{adj(A)}{|A|} \tag{Α΄.7}$$

όπου $adj(A)$ είναι ο $n \times n$ πίνακας του οποίου το στοιχείο $(i,j)$ είναι ίσο με $(-1)^{(i+j)}D_{ji}$ **προσοχή στην αντιστροφή της σειράς των δεικτών i,j.**. Ο πίνακας $adj(A)$ ονομάζεται *συζυγής* του $A$.

## Α΄.3 Γραμμική Ανεξαρτησία και Βαθμός

Τα διανύσματα $x_1, x_2, ..., x_n$ λέγεται ότι είναι *γραμμικώς εξαρτημένα* αν και μόνο αν υπάρχουν βαθμωτές σταθερές $c_1, ..., c_n$, όχι όλες μηδενικές, τέτοιες ώστε:

$$c_1x_1 + c_2x_2 + ... + c_nx_n = 0 \tag{Α΄.8}$$

Αν δεν υπάρχουν τέτοιες σταθερές, τότε τα διανύσματα $x_i$ λέγεται οτι είναι γραμμικώς *ανεξάρτητα*. Με άλλα λόγια, τα $x_1, ..., x_n$ είναι ανεξάρτητα αν κανένα απ'αυτά δεν μπορεί να εκφραστεί σαν γραμμικός συνδιασμός των υπολοίπων.

Αν $A \in \mathbb{R}^{n\times m}$, τότε () ή $rank(A)$ είναι ο μεγαλύτερος αριθμός ανεξαρτήτων σειρών (ή στηλών) του $A$. Επίσης,

- $rank(A) \leq min(n,m)$. Εάν ισχύει ισότητα, τότε ο $A$ λέγεται ότι είναι "πλήρους βαθμού".

- $rank(AB) \leq min(rank(A), rank(B))$
- $rank(AB) = rank(A)$ αν υπάρχει ο $B^{-1}$
- $rank(AB) = rank(B)$ αν υπάρχει ο $A^{-1}$

Η έννοια του βαθμού πίνακα είναι στενά συνδεδεμένη με αυτή του μηδενοχώρου ενός πίνακα. Ο *μηδενοχώρος* του $A \in \mathbb{R}^{n\times m}$ (συμβ. $\mathcal{N}A$) είναι το σύνολο των διανυσμάτων $x \in \mathbb{R}^m$ για τα οποία $Ax = 0$.

Το παρακάτω θεώρημα αφορά στην ύπαρξη λύσεων για γραμμικές αλγεβρικές εξισώσεις

**Θεώρημα 3.1**

$Ax = b$ έχει λύση ώς προς το $x$, ή υπάρχει $p \in \mathcal{N}(A^T)$ τέτοιο ώστε $\langle p, b\rangle \neq 0$ (δηλ. το $b$ δεν είναι ορθογώνιο στον μηδενοχώρο του $A^T$)

Απόδειξη: $Ax = b$ έχει λύση αν το $b$ βρίσκεται στο χώρο στηλών του $A$, $Range(A)$. Αλλά $b \in Range(A)$ αν και μόνο αν $b \perp \mathcal{N}(A^T)$. Η τελευταία σχέση συνεπάγεται ότι αν η $Ax = b$ δεν έχει λύση, θα υπάρχει στοιχείο του $\mathcal{N}(A^T)$ - έστω το $p$ - το οποίο δεν είναι ορθογώνιο με το $b$, δηλ. $\langle p, b\rangle \neq 0$. $\square$.

## Α΄.4 Ιδιοτιμές και Ιδιοδιανύσματα

Ένα διάνυσμα $v \neq 0$ λέγεται *ιδιοδιάνυσμα* (eigenvector) του πίνακα $A$ αν

$$Av = \lambda v \tag{Α΄.9}$$

για κάποια βαθμωτή μεταβλητή, $\lambda$, δηλ. αν το $v$ είναι παράλληλο με τό $Av$. Για να ισχύει η τελευταία εξίσωση πρέπει να έχουμε

$$(\lambda I - A)v = 0 \tag{Α΄.10}$$

Επειδή υποθέσαμε ότι $v \neq 0$, συμπεραίνουμε ότι ο πίνακας

$$\lambda I - A = \begin{bmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{bmatrix} \tag{Α΄.11}$$

έχει μηδενική ορίζουσα. Προσέξτε ότι άν το $v$ είναι ιδιοδιάνυσμα του $A$ τότε και το $\alpha v$ είναι ιδιοδιάνυσμα του $A$ για οποιοδήποτε $\alpha$ πραγματικό ή μιγαδικό.

Η ορίζουσα του $sI - A$ είναι ένα πολυώνυμο $n$-οστού βαθμού ως προς την ( μιγαδική μεταβλητή) $s$, και λέγεται *χαρακτηριστικό πολυώνυμο* του $A$:

$$p_A(s) = det(sI - A) = s^n + p_{n-1}s^{n-1} + ... + p_1 s + p_0 \tag{Α΄.12}$$

όπου $p_0, ..., p_{n-1}$ είναι (πραγματικοί) συντελεστές που ορίζονται με βάση τα στοιχεία του $A$. (**Ερώτηση: Γιατί ο συντελεστής του $s^n$ είναι πάντα $1$?**)

Το χαρακτηριστικό πολυώνυμο ενός πίνακα $n \times n$ $A$, $det(sI - A) = 0$ θα έχει $n$ ρίζες με τιμές $s = \lambda_1, ..., s = \lambda_n$ οι οποίες λέγονται *χαρακτηριστικές τιμές* ή *ιδιοτιμές* του $A$. Δηλαδή, μπορούμε να ξαναγράψουμε την (Α΄.12) ως εξής:

$$p_A(s) = (s - \lambda_1)(s - \lambda_2) \cdots (s - \lambda_n) \qquad \text{(Α΄.13)}$$

Οι ιδιοτιμές μπορεί να είναι πραγματικές ή μιγαδικές. Αν ο $A$ είναι πραγματικός, οι μιγαδικές ιδιοτιμές του εμφανίζονται πάντα σε συζυγή ζεύγη της μορφής $a + ib$, $a - ib$.

### Πολλαπλότητα Ιδιοτιμής

Γιά κάθε ιδιοτιμή $\lambda_i$, μπορούμε να βρούμε διάνυσμα $v_i$ που ικανοποιεί

$$Av_i = \lambda_i v_i \qquad \text{(Α΄.14)}$$

Αν όλες οι ιδιοτιμές του $A$'s είναι διακριτές (δηλ. όλες έχουν διαφορετικές τιμές) τότε ο $A$ έχει $n$ ανεξάρτητα ιδιοδιανύσματα. Είναι δυνατό όμως ο $A$ (το χαρακτηριστικό του πολυώνυμο) να έχει επαναλαμβανόμενες ιδιοτιμές (ρίζες), (π.χ. αν $p_A(s) = (s - 1)^3(s + 5)$, τότε οι ιδιοτιμές του $A$ είναι $1, 1, 1$ και $-5$). Ο αριθμός των φορών τις οποίες η ιδιοτιμή $\lambda$ εμφανίζεται σαν ρίζα του χαρ. πολυωνύμου λέγεται *αλγεβρική πολλαπλότητα*, $n_a(\lambda)$, της ιδιοτιμής (π.χ. η ιδιοτιμή $\lambda = 1$ έχει αλγεβρική πολλαπλότητα 3 στο προηγούμενο παράδειγμα). Εάν η $\lambda$ είναι επαναλαμβανόμενη τότε *είναι δυνατό να μην υπάρχουν παραπάνω από 1 ανεξάρτητα ιδιοδιανύσματα* που ικανοποιούν την (Α΄.14) για $\lambda_i = \lambda$. Ο αριθμός των ανεξάρτητων ιδιοδιανυσμάτων λέγεται *γεωμετρική πολλαπλότητα* της ιδιοτιμής $\lambda$ (συμβ. $n_g(\lambda)$). Προφανώς, $n_g(\lambda) \leq n_a(\lambda)$.

## Α΄.5 Μετασχηματισμοί ομοιότητας και Διαγωνιοποίηση

Οποιοδήποτε ζεύγος πινάκων $A$ και $B$ που σχετίζονται με το μετασχηματισμό

$$B = TAT^{-1} \qquad \text{(Α΄.15)}$$

όπου $T$ μη-ιδιάζων πίνακας ($det(T) \neq 0$)), λεγονται *όμοιοι* μεταξύ τους. Οι μετασχηματισμοί ομοιότητας δεν αλλάζουν το χαρακτηριστικό πολυώνυμο (και άρα τις ιδιοτιμές και την ορίζουσα) ενός πίνακα. (**Να αποδειχθεί.**)

Έστω $v_1, ..., v_n$ τα (ανεξάρτητα) ιδιοδιανύσματα ενός πίνακα $A$, του οποίου οι ιδιοτιμές είναι $\lambda_1, ..., \lambda_n$, όλες διακριτές μεταξύ τους. Αν τοποθετήσουμε τα ιδιοδιανύσματα σε έναν πίνακα $V$:

$$V = \Big[ v_1 \big| v_2 \big| \cdots \big| v_n \Big] \qquad \text{(Α΄.16)}$$

και τις ιδιοτιμές σε ένα διαγώνιο πίνακα $\Lambda$:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \tag{Α΄.17}$$

τότε η Εξ.Α΄.14 δίνει:

$$V\Lambda = AV \tag{Α΄.18}$$

Επειδή τα ιδιοδιανύσματα είναι ανεξάρτητα, μπορούμε να γράψουμε:

$$A = V\Lambda V^{-1} \tag{Α΄.19}$$

το οποίο μας οδηγεί στο εξής σημαντικό γεγονός: Κάθε πίνακας $A$ με διακριτές ιδιοτιμές μπορεί να διαγωνιοποιηθεί μέσω ενός μετασχηματισμού ομοιότητας (συγκεκριμένα αυτού που αποτελείται απο τα ιδιοδιανύσματα του $A$'). Με άλλα λόγια, "ο $A$ γίνεται διαγώνιος όταν εκφραστεί στη βάση των ιδιοδιανυσμάτων του".

### Α΄.5.1 Η μορφή Jordan

Τι γίνεται όμως αν ο $A$ έχει επαναλαμβανόμενες ιδιοτιμές; Σε αυτή την περίπτωση δεν είναι πάντα εφικτή η διαγωνιοποίηση του $A$. Είναι όμως πάντα εφικτό να μετασχηματίσουμε τον $A$ σε μια μορφή η οποία είναι "όσο διαγώνια γίνεται". Εδώ παρουσιάζουμε τα βήματα για την κατασκευή αυτής της μορφής.

Ας υποθέσουμε ότι ο $A$ εχει διαστάσεις $n \times n$ και μπορεί να έχει επαναλαμβανόμενες πραγματικές ιδιοτιμές, αλλά όχι επαναλαμβανόμενες μιγαδικές ιδιοτιμές. Αν ο $A$ έχει $q$ ανεξάρτητα ιδιοδιανύσματα, τότε θα είναι όμοιος προς έναν πίνακα $J$, επίσης $n \times n$, ο οποίος αποτελείται από $q$ υποπίνακες ("μπλοκ"):

$$J = M^{-1}AM = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & \ddots & 0 & 0 \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & J_q \end{bmatrix} \tag{Α΄.20}$$

(όπου κάθε "0" αντιπροσωπεύει ένα μπλοκ απο μηδενικά, με κατάλληλες διαστάσεις). Κάθε "μπλοκ Jordan", $J_i$, $i = 1, ..., q$, είναι ένας μικρότερος πίνακας για τον οποίο ισχύει ένα απο τα δύο:

- ή $J_i$ έχει μία και μοναδική πραγματική ιδιοτιμή $\lambda_i$ (μπορεί αυτή να είναι επαναλαμβανόμενη) και μόνο ένα ιδιοδιάνυσμα, ή
- ο $J_i$ έχει ένα ζεύγος συζυγών μιγαδικών ιδιοτιμών και ένα αντίστοιχο ζεύγος συζυγών μιγαδικών ιδιοδιανυσμάτων.

Η δομή των μπλοκ Jordan είναι μια απο τις παρακάτω 3:

- I. $\lambda_i$ πραγματική και μοναδική, τότε:

$$J_i = \lambda_i \tag{Α΄.21}$$

  (δηλ. ο $J_i$ είναι απλά η ιδιοτιμή).
- II. $\lambda_i$ πραγματική και επαναλαμβανόμενη. Τότε

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots\cdots & 0 \\ 0 & \lambda_i & 1 & 0 \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots\cdots\cdots & & \lambda_i & 1 \\ 0 & \cdots\cdots\cdots & & 0 & \lambda_i \end{bmatrix} \tag{Α΄.22}$$

  Παρατηρούμε ότι τέτοιας μορφής μπλοκ Notice that such Jordan έχουν *μόνο ένα* (ανεξάρτητο) ιδιοδιάνυσμα, δηλ. μόνο ένα ανεξάρτητο διάνυσμα $v$ που ικανοποιεί τη σχέση $J_i v = \lambda_i v$.
- III. $\lambda_i = a + ib$ μιγαδική ιδιοτιμή. Τότε η $a - ib$ είναι και αυτή ιδιοτιμή του $A$. Σε αυτή την περίπτωση,

$$J_i = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \tag{Α΄.23}$$

  Τέλος, οι παρακάτω “κανόνες” καθορίζουν το μέγεθος και πλήθος των μπλοκ Jordan:

- 1. Ο συνολικός αριθμός εμφανίσεων της $\lambda_i$ στη διαγώνιο του $J$ *είναι ίσος* με την αριθμητική πολλαπλότητα της $\lambda_i$.
- 2. Ο αριθμός των μπλοκ $J_i$ στα οποία η $\lambda_i$ εμφανίζεται στη διαγώνιο, ισούται με τη γεωμετρική πολλαπλότητα της $\lambda_i$.

  Παράδειγμα: έστω

$$A_1 = \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \quad A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \tag{Α΄.24}$$

Αρχικά, δείξτε (ή πείστε τον εαυτό σας) ότι τα χαρακτηριστικά πολυώνυμα των $A_1$ and $A_2$ είναι ίδια:

$$p_{A_1}(s) = s^3(s+3) \tag{Α΄.25}$$
$$p_{A_2}(s) = s^3(s+3). \tag{Α΄.26}$$

Οπότε, οι ιδιοτιμές και των δυο πινάκων $A_1, A_2$ είναι $0, 0, 0, -3$. Η ιδιοτιμή $s = 0$ είναι επαναλαμβανόμενη, με αλγεβρική πολλαπλότητα 3, ενώ η $s = -3$ είναι μη-επαναλαμβανόμενη με αλγεβρική πολλαπλότητα 1. Αλλά, στην περίπτωση του $A_1$, το μόνο ιδιοδιάνυσμα που αντιστοιχεί στην $s = 0$ είναι το $(1, 0, 0, 0)^T$. Άρα η μορφή Jordan του $A_1$'s θα έχει μόνο ένα μπλοκ που θα αντιστοιχεί στην $s = 0$ (η γεωμ. πολλαπλότητά της είναι 1). Η ιδιοτιμή $s = -3$ είναι μη-επαναλαμβανόμενη, άρα το μπλοκ Jordan που αντιστοιχεί σε αυτή θα είναι $1 \times 1$ (δηλαδή απλά η ίδια η ιδιοτιμή). Συμπεραίνουμε ότι η μορφή Jordan του $A_1$ είναι:

$$J_1 = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -3 \end{array}\right] \tag{Α΄.27}$$

Στη συνέχεια ας ελέγξουμε τον $A_2$. Εχει το ίδιο χαρακτηριστικό πολυώνυμο με τον $A_1$, αλλά στη περίπτωση του $A_2$'s η ιδιοτιμή $s = 0$ έχει δύο ανεξάρτητα ιδιοδιανύσματα, $(1, 0, 0, 0)^T$ and $(0, 1, 0, 0)$ (μπορείτε να ελέγξετε ότι επαληθεύουν τη σχέση $A_2 v = 0 \cdot v$. Αυτό μας λεει ότι η μορφή Jordan του $A_2$ (ας την ονομάσουμε $J_2$) θα περιέχει 2 μπλοκ που αντιστοιχούν στην ιδιοτιμή $s = 0$:

$$J_2 = \left[\begin{array}{cc|c|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -3 \end{array}\right] \tag{Α΄.28}$$

δηλαδή ένα $2 \times 2$ μπλοκ για τα 2 "αντίγραφα" της ιδιοτιμής $s = 0$, ένα $1 \times 1$ μπλοκ για την 3η ιδιοτιμή στο $s = 0$, και ένα $1 \times 1$ για την ιδιοτιμή $s = -3$.

## Α΄.6 Ειδικές κατηγορίες Πινάκων

Για πραγματικούς πίνακες ισχύουν:

- Ο $A$ είναι ορθογώνιος αν $A^T = A^{-1}$.
- Ο $A$ είναι συμμετρικός αν $A = A^T$. Οι συμμετρικοί πίνακες έχουν πραγματικές ιδιοτιμές.
- Ο $A$ είναι αντι-συμμετρικός αν $A = -A^T$.
- Ο $A$ είναι άνω (κάτω) τριγωνικός άν όλα τα στοιχεία του κάτω (πάνω) από τη διαγώνιο είναι μηδενικά.

Αν ένας συμμετρικός πίνακας $M$ ($n \times n$) ικανοποιεί

$$x^T M x \geq 0 \quad \forall x \in \mathbb{R}^n$$

τότε ο λέγεται **θετικά ημι-ορισμένος**. Λέγεται **θετικά ορισμένος** αν η ισχύει αυστηρά η ανισότητα. Εαν η ανισότητα αντιστραφεί

$$x^T M x \leq 0 \quad \forall x \in \mathbb{R}^n$$

τότε μιλάμε για πίνακα **αρνητικά (ημι-)ορισμένο**. **Προσοχή: Ο θετικός (αρνητικός) ορισιμότητα είναι ιδιότητα μόνο των συμμετρικών πινάκων**.

Για $Q = Q^T > 0$, $Q \in \mathbb{R}^{n \times n}$, τα σημεία $x \in \mathbb{R}^n$ τα οποία ικανοποιούν την $y = x^T Q x$ για δεδομένη (μη-αρνητική) τιμή του $y$, ορίζουν ένα ελλειψοειδές στο $\mathbb{R}^n$. Για παράδειγμα, εαν

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

τότε η $y = x^T Q x = x_1^2 + x_2^2$, ορίζει κύκλο ακτίνας $\sqrt{y}$. Γενικά, η συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = x^T Q x$, όπου ο $Q$ είναι θ.ο., ορίζει μια οικογένεια από ελειψοειδή στο $\mathbb{R}^n$.

### Α΄.6.1 Τέστ για θετική ορισιμότητα

Γνωρίζουμε ότι οποιοσδήποτε συμμετρικός πίνακας $Q = Q^T$ έχει πραγματικές ιδιοτιμές και μπορεί να διαγωνιοποιηθεί από ορθογώνιο πίνακα, δηλ. υπάρχει $\Theta$ με $\Theta^T = \Theta^{-1}$ ώστε

$$\Theta Q \Theta^T = D; \quad D \triangleq diag(\lambda_1, \lambda_2, ..., \lambda_n)$$

Ο $Q$ είναι θετικά ορισμένος αν και μόνο αν $x^T Q x > 0$ η, ισοδύναμα,

$$x^T \Theta^T D \Theta x = y^T D y = \sum \lambda_i y_i^2 > 0$$

όπου θέσαμε $y = \Theta x$. Προφανώς, το τελευταίο άθροισμα είναι θετικό αν και μονο αν όλα τα $\lambda_i > 0$. Επομένως, δείξαμε ότι ο $Q$ είναι θετικά ορισμένος αν και μόνο αν όλες οι ιδιοτιμές του είναι θετικές.

Επίσης, ο $Q = Q^T$ είναι θετικά ορισμένος αν όλες οι κύριες ελλάσονες του

$$P_1 = q_{11}, P_2 = det \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}, ...P_n = |Q|$$

είναι **θετικές**. (Σημ.: Ο $Q = Q^T$ είναι **αρνητικά ορισμένος** αν και μόνο αν όλες οι ιδιοτιμές του είναι αρνητικές ή αν οι κύριες ελλάσονές του έχουν **εναλλασσόμενα πρόσημα** με $P_1 < 0$, $P_2 > 0$, $P_3 < 0$,... κλπ.)

## Α΄.7 Σημαντικές Ιδιότητες Πινάκων

- Μη-αντιμεταθετικότητα: Γενικά, για ένα ζεύγος τετράγωνων πινάκων $A, B$, $AB \neq BA$
- Πολλαπλάσια του μοναδιαίου πίνακα αντιμετατίθονται με όλους τους πίνακες (συμβατών διαστάσεων).
- (Θεώρημα Caley-Hamilton): Ένας τετράγωνος πίνακας A ικανοποιεί το χαρακτηριστικό του πολυώνυμο,
  δηλ. αν $p_A(s) = s^n + p_{n-1}s^{n-2} + \cdots + p_1 s + p_0$, τότε $p_A(A) = 0$.
- Το άθροισμα των διαγώνιων στοιχείων ενός πίνακα $A$ (γνωστό ως το *ίχνος (trace)* του $A$ ή $tr(A)$) ισούται με το άθροισμα των ιδιοτιμών του $A$: $tr(A) = \sum_1^n a_{ii}$ (**Να αποδειχθεί**).
- Αν οι $A, B$ έχουν συμβατές διαστάσεις, τότε $tr(AB) = tr(BA)$ (**Να αποδειχθεί**).
- Η ορίζουσα του $A$ ισούται με το γινόμενο των ιδιοτιμών του $A$: $|A| = \prod_1^n \lambda_i(A)$ (**Να αποδειχθεί**).
- Αν ο $A$ είναι άνω (κάτω) τριγωνικός, τότε τα διαγώνια στοιχεία του $A$ *είναι* οι ιδιοτιμές του $A$ (**Να αποδειχθεί**).

# Β′

## Παραγώγιση συναρτήσεων διανυσματικής μεταβλητής

Έστω $f(x) : \mathbb{R}^n \to \mathbb{R}$ παραγωγίσιμη. Θα επιλέξουμε να γράφουμε την παράγωγο της $f$ σε μορφή διανύσματος-στήλης, δηλαδή:

$$\frac{df}{dx} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \tag{Β′.1}$$

Για $f(x) : \mathbb{R}^n \to \mathbb{R}^m$, παραγωγίσιμη, με

$$f(x) = [f_1(x), f_2(x), ..., f_m(x)]^T$$

η παράγωγος $df/dx$ ορίζεται ως εξής:

$$\frac{df}{dx} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Πρακτικά, ο υπολογισμός παραγώγων διανυσματικών συναρτήσεων με διανυσματικές μεταβλητές γίνεται με βάση κανόνες ίδιους με αυτούς που διέπουν την παραγώγιση συναρτήσεων μίας μεταβλητής, λαμβάνοντας όμως υπ'όψιν την διανυσματική φύση των διαφόρων όρων. Για παράδειγμα, εαν $f(x) = a^T x$, όπου $x, a \in \mathbb{R}^n$, τότε

$$df/dx = \frac{d(a^T x)}{dx} = a \tag{Β′.2}$$

(συγκρίνετε την τελευταία εξίσωση με την παράγωγο της $f(x) = ax$, όπου $x, a \in \mathbb{R}$.) Το τελευταίο αποδεικνύεται εύκολα αν λάβουμε υπ'όψι ότι $f(x) = a_1x_1 + a_2x_2 + ... + a_nx_n$, οπότε η στήλη $i$ της δεξιάς πλευράς της (Β′.1) είναι απλά το $a_i$.

Επίσης, έστω $f(x) = x^T Qx$, $Q = Q^T > 0$, $x \in \mathbb{R}^n$.

$$\frac{df}{dx} = 2Qx$$

(συγκρίνετε την τελευταία εξίσωση με την παράγωγο της $f(x) = qx^2$, όπου $x, q \in \mathbb{R}$.) Για να αποδειχθεί η τελευταία εξίσωση, αρκεί να υπολογίσουμε τις μερικές παραγώγους $df/dx_i$, $i = 1, 2, ..., n$ (Ασκηση). Εναλλακτικά, μπορούμε να εφαρμόσουμε τον κανόνα για την παραγώγιση γινομένου, λαμβάνοντας το άθροισμα που προκύπτει από την παραγώγιση της $f$ ως προς κάθε μεταβλητή χωριστά. Για παράδειγμα, χρησιμοποιόντας την (Β′.2), έχουμε

$$\frac{d(y^T Qx)}{dx} = Q^T y$$

και, εφ'οσον λόγω της συμμετρίας του $Q$ έχουμε $x^T Qy = (x^T Qy)^T = y^T Q^T x = y^T Qx$,

$$\frac{d(x^T Qy)}{dx} = \frac{d(y^T Q^T x)}{dx} = Qy$$

και τελικά,

$$\frac{d(x^T Qx)}{dx} = \frac{d(x^T Qx)}{dx} + \frac{d(x^T Q^T x)}{dx} = Qx + Q^T x = 2Qx$$

όπου στην τελευταία εξίσωση παραγωγίσαμε την $x^T Qx$ ως προς τις $x$ που εμφανίζονται αριστερά και δεξιά του $Q$, λαμβάνοντας υπόψη μας ότι η ποσότητα $x^T Qx$ είναι βαθμωτή (scalar) άρα η αναστροφή αυτής δεν αλλάζει το αποτέλεσμα.