



ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Σχολη ΗΜ&ΜΥ, ΕΜΠ
Καθηγητης Πετρος Μαραγκος

Στοιχεια Γραμμικης Αλγεβρας,
Ελαχιστα Τετραγωνα, SVD,
Θετικα-ορισμενοι Πινακες

Ειδικες Μορφες Πινακων

- Ένας **διαγώνιος** πίνακας είναι τετραγωνικός πίνακας της μορφής:

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

- Όταν όλα τα στοιχεία της διαγωνίου είναι ίσα με μονάδα ο πίνακας αυτός καλείται **μοναδιαίος**.

Τριγωνικός Πίνακας

- **Ανω τριγωνικός** πίνακας είναι ο πίνακας όπου όλα τα στοιχεία του κάτω από τη διαγώνιο είναι μηδενικά

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

- Αντίθετα ο πίνακας του οποίου όλα τα στοιχεία πάνω από τη διαγώνιο είναι μηδενικά ονομάζεται κάτω τριγωνικός.

Πίνακας Toeplitz

- Ένας **Toeplitz** πίνακας \mathbf{T} $N \times N$ είναι ένας πίνακας που έχει σταθερά όλα τα στοιχεία του στην κύρια διαγώνιο και στις υποδιαγωνίους. Δηλαδή το στοιχείο $t(m,n)$ εξαρτάται μόνο από τη διαφορά $m-n$, $t(m,n)=t_{m-n}$.

$$\mathbf{T} = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-N+1} \\ t_1 & t_0 & t_{-1} & t_{-N+2} \\ t_2 & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & t_{-1} \\ t_{N-1} & \cdots & t_2 & t_1 & t_0 \end{bmatrix}$$

Επιπλέον καθορίζεται πλήρως από $(2N-1)$ στοιχεία $\{t_k, -N+1 \leq k \leq N-1\}$

Κυκλικός Πίνακας

- Ένας πίνακας C είναι **κυκλικός** εάν κάθε γραμμή (στήλη) του πίνακα αποτελεί μια κυκλική μετατόπιση της προηγούμενης γραμμής (στήλης)

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{N-1} \\ c_{N-1} & c_0 & c_1 & & c_{N-2} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ c_2 & & & & c_1 \\ c_1 & c_2 & \cdots & c_{N-1} & c_0 \end{bmatrix}$$

- Ένας κυκλικός πίνακας είναι Toeplitz και επιπλέον ισχύει

$$c(m, n) = c((m - n) \text{ modulo } N)$$

Γραμμική Συνελίξη και Πίνακας Toeplitz

- Η έξοδος ενός ΓΧΑ συστήματος με απόκριση $h(n) = n$, $-1 \leq n \leq 1$ και είσοδο $x(n)$ που είναι μηδενική εκτός του $0 \leq n \leq 4$ δίνεται από:

$$y(n) = h(n) \circledast x(n) = \sum_{k=0}^4 h(n-k)x(k)$$

- Η συνέλιξη αυτή μπορεί να γραφεί ως γινόμενο Toeplitz πίνακα επί διάνυσμα

$$\begin{bmatrix} y(-1) \\ y(0) \\ y(1) \\ y(2) \\ y(3) \\ y(4) \\ y(5) \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \\ x(4) \end{bmatrix}$$

Κυκλική Συνέλιξη και Κυκλικός Πίνακας

- Εάν οι υπό συνέλιξη ακολουθίες είναι περιοδικές τότε η συνέλιξη τους είναι επίσης περιοδική και μπορεί να αναπαρασταθεί ως:

$$y(n) = \sum_{k=0}^{N-1} h(n-k)x(k), \quad 0 \leq n \leq N-1$$

- Παράδειγμα κυκλικής συνέλιξης ($N=4$) ως γινόμενο κυκλικού πίνακα επί διάνυσμα ($h[n] = n+3 \text{ modulo } 4$).

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \end{bmatrix} = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 0 & 3 & 2 & 1 \\ 1 & 0 & 3 & 2 \\ 2 & 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \end{bmatrix}$$

Πίνακας Hankel

- **Hankel** ονομάζεται ο πίνακας που έχει ίδια στοιχεία στις διαγωνίους οι οποίες είναι κάθετες στην κύρια διαγώνιο

$$a_{ij} = a_{i+1,j-1} \quad ; \quad \text{for all } i < n \text{ and } j \leq n$$

- Παράδειγμα Hankel πίνακα

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 5 & 7 & 4 \\ 5 & 7 & 4 & 2 \\ 7 & 4 & 2 & 1 \end{bmatrix}$$

Πινακας Ανταλλαγης (Exchange)

$$\mathbf{J} = \begin{bmatrix} 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 1 & 0 & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ 1 & \dots & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{J}^2 = \mathbf{I}$$

$$\mathbf{J} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} v_n \\ v_{n-1} \\ \vdots \\ v_1 \end{bmatrix}$$

Ιδιοτητες Πινακα Ανταλλαγης

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \mathbf{J}^T \mathbf{A} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{bmatrix}$$

$$\mathbf{A} \mathbf{J} = \begin{bmatrix} a_{13} & a_{12} & a_{11} \\ a_{23} & a_{22} & a_{21} \\ a_{33} & a_{32} & a_{31} \end{bmatrix} \quad \mathbf{J}^T \mathbf{A} \mathbf{J} = \begin{bmatrix} a_{33} & a_{32} & a_{31} \\ a_{23} & a_{22} & a_{21} \\ a_{13} & a_{12} & a_{11} \end{bmatrix}$$

Περ-συμμετρικός Πίνακας

$$a_{ij} = a_{n-j+1, n-i+1}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 2 & 2 & 4 & 5 \\ 4 & 4 & 2 & 3 \\ 6 & 4 & 2 & 1 \end{bmatrix}$$

Κεντρο-Συμμετρικός Πίνακας (= Συμμετρικός & Περ-συμμετρικός)

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 & 6 \\ 3 & 2 & 4 & 5 \\ 5 & 4 & 2 & 3 \\ 6 & 5 & 3 & 1 \end{bmatrix}$$

Συμμετρικός Toeplitz (πραγματικός)

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 & 7 \\ 3 & 1 & 3 & 5 \\ 5 & 3 & 1 & 3 \\ 7 & 5 & 3 & 1 \end{bmatrix}$$

$$\mathbf{J}^T \mathbf{A} \mathbf{J} = \mathbf{A}$$

The Relationship between the Symmetries of a Matrix and Its Inverse

Matrix

Symmetric

Hermitian

Persymmetric

Centrosymmetric

Toeplitz

Hankel

Triangular

Inverse

Symmetric

Hermitian

Persymmetric

Centrosymmetric

Persymmetric

Symmetric

Triangular

Στοιχεία Γραμμικής Αλγεβρας

Ref: G. Strang, *Linear Algebra and Its Applications*, 1986.

- Matrix Factorizations (LDU, Spectral, SVD)
- SVD and Least Squares
- Hermitian and Positive-definite matrices
- Fundamental Spaces
- Ellipsoids and Positive-definite Matrices
- Generalized Eigenvalue Problem
- Min-Max Principle and Rayleigh Quotient

Θεμελιώδεις Χωροι Πινακων και Ελαχιστα Τετραγωνα

Θεμελιώδες Θεώρημα Γραμμικής Αλγεβρας

Συστημα m Γραμμικων Εξισωσεων με n αγνωστους: $\mathbf{Ax} = \mathbf{b}$

\mathbf{A} παριστανει γραμμικο μετ/σμο απο τον χωρο \mathbb{R}^n στον χωρο \mathbb{R}^m

$$r = \text{rank}(\mathbf{A}) \leq \min(m, n)$$

$R(\mathbf{A}) =$ χωρος στηλων, $R(\mathbf{A}^T) =$ χωρος γραμμων, $N(\cdot) =$ μηδενοχωρος

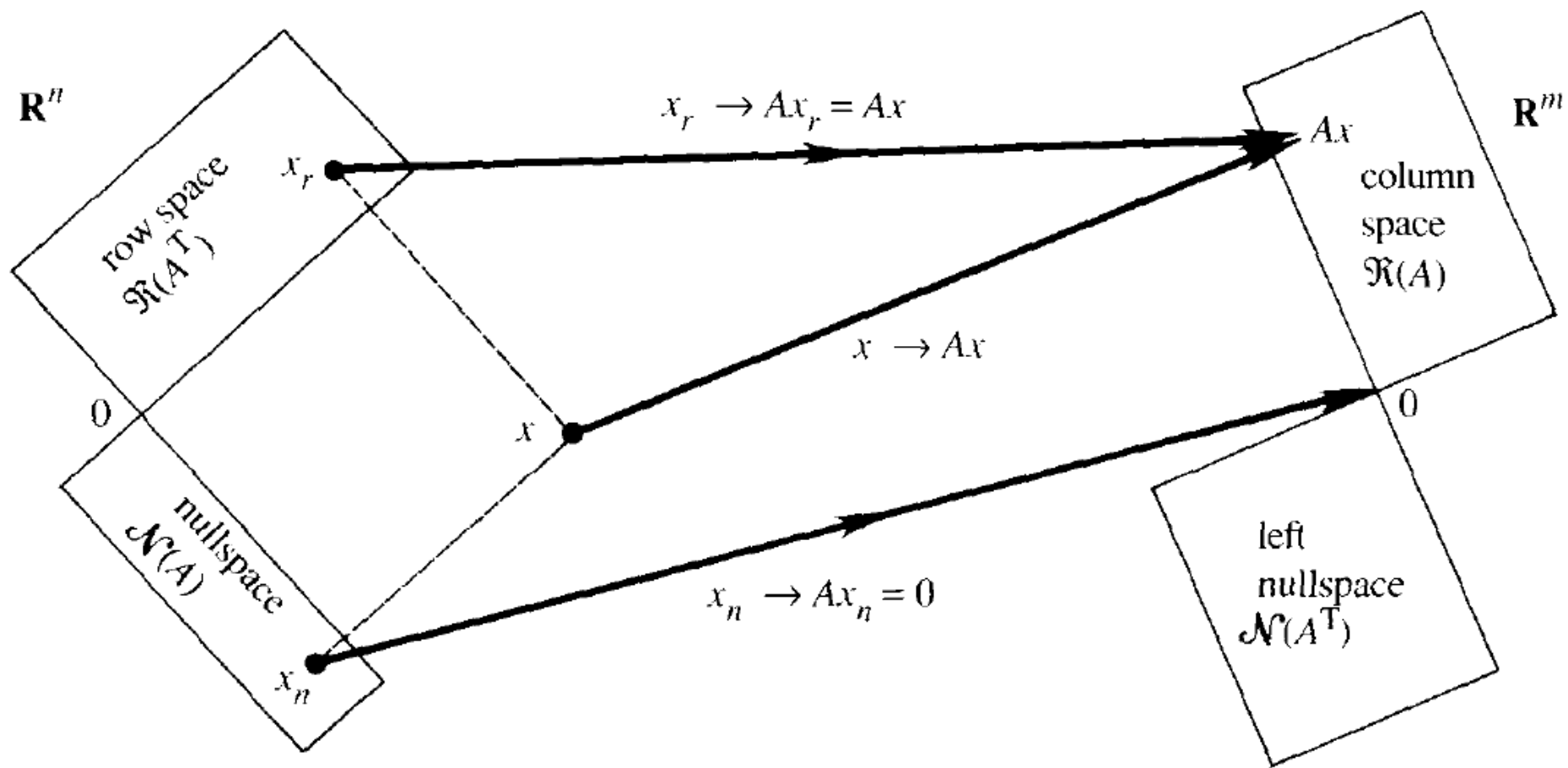
THEOREM 1 (FUNDAMENTAL THEOREM OF LINEAR ALGEBRA)

(a) *The dimensions of the four fundamental spaces of a real $m \times n$ matrix \mathbf{A} are:*

$$\begin{aligned} \text{column space} & : \dim \mathcal{R}(\mathbf{A}) = r \\ \text{row space} & : \dim \mathcal{R}(\mathbf{A}^T) = r \\ \text{null space} & : \dim \mathcal{N}(\mathbf{A}) = n - r \\ \text{left null space} & : \dim \mathcal{N}(\mathbf{A}^T) = m - r \end{aligned} \tag{1}$$

(b) *In \mathbb{R}^m , the orthogonal complement of the column space of \mathbf{A} is the null space of its transpose. In \mathbb{R}^n , the orthogonal complement of the null space of \mathbf{A} is the row space of its transpose.*

Δράση ενός πίνακα $m \times n$ \mathbf{A} για λύση m εξισώσεων με n αγνώστους



Επίλυση Γραμμικών Εξισώσεων και Ελαχιστα Τετραγωνα

Case I (Full Rank): $r = \min(m, n)$:

If $r = m \leq n$ (independent rows), then there exists *at least one solution*.

If $r = n < m$ (independent columns), we have *at most one solution*.

$\mathbf{b} \notin R(\mathbf{A}) \Rightarrow$ καμία λύση \rightarrow

προσεγγιστική λύση με ελαχιστα τετραγωνα

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \underbrace{\|\mathbf{Ax} - \mathbf{b}\|}_{\mathbf{e}}$$

$$\frac{\partial \mathbf{e}^T \mathbf{e}}{\partial \mathbf{x}} = 0 \Leftrightarrow \mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b} \Leftrightarrow \hat{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$$

$$\hat{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{b}, \quad \mathbf{A}^\dagger \triangleq (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

\mathbf{A}^\dagger is the *Moore-Penrose pseudo-inverse*

Case II (Low Rank): $r < \min(m, n)$:

προσεγγιστική λύση ελαχιστων τετραγωνων με ελαχιστη νορμα

$$\mathbf{x}^+ = \mathbf{A}^+ \mathbf{b}, \quad \mathbf{A}^+ = \text{γενικός ψευδο-αντιστροφος απο SVD}$$

Παραγοντοποιήσεις Πινάκων

- Cholesky
- Spectral Factorization
- SVD

Τριγωνική Παραγοντοποίηση: Με την απαλοιφή Gauss, υποθέτοντας ότι δεν απαιτούνται εναλλαγές γραμμών, κάθε πίνακας \mathbf{A} μπορεί να παραγοντοποιηθεί ως $\mathbf{A} = \mathbf{LDU}$, όπου \mathbf{L} και \mathbf{U} είναι κάτω και άνω τριγωνικοί πίνακες αντίστοιχα με διαγώνιους όρους ίσους με μονάδα, και \mathbf{D} είναι ο διαγώνιος πίνακας με τους πιλοτικούς όρους. Εάν ο \mathbf{A} είναι αντιστρέψιμος, η παραγοντοποίηση αυτή είναι μοναδική. Εάν ο \mathbf{A} είναι Ερμιτιανός-Hermitian, τότε $\mathbf{U} = \mathbf{L}^H$ και προκύπτει ότι

$$\mathbf{A} = \mathbf{LDL}^H \quad (1)$$

Cholesky διάσπαση: Εάν ο \mathbf{A} είναι Hermitian και θετικά ημι-ορισμένος, τότε ο \mathbf{D} έχει μη αρνητικούς διαγώνιους όρους. Επομένως, η τριγωνική παραγοντοποίηση γίνεται

$$\mathbf{A} = \mathbf{L}\sqrt{\mathbf{D}}\sqrt{\mathbf{D}}\mathbf{L}^H = \mathbf{L}\sqrt{\mathbf{D}}(\mathbf{L}\sqrt{\mathbf{D}})^H \quad (2)$$

που είναι το γινόμενο του κάτω τριγωνικού πίνακα με τον Hermitian ανάστροφο του. Συχνά η ανωτέρω καλείται διάσπαση Cholesky και γράφεται σε ποιο συμπαγή μορφή όπως στο (1).

Φασματική Διάσπαση: Κάθε $n \times n$ πίνακας \mathbf{A} που διαθέτει n γραμμικά ανεξάρτητα ιδιοδιανύσματα δέχεται μια *διάσπαση ιδιοτιμών*

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} \quad (3)$$

όπου \mathbf{V} περιέχει ως στήλες τα ιδιοδιανύσματα $\mathbf{v}_1, \dots, \mathbf{v}_n$ και $\mathbf{\Lambda}$ είναι ένας διαγώνιος πίνακας που περιέχει τις ιδιοτιμές $\lambda_1, \dots, \lambda_n$. Αυτή η παραγοντοποίηση καλείται επίσης *διαγωνοποίηση* του \mathbf{A} .

Κανονικοί (Normal) πίνακες \mathbf{A} , δηλ. πίνακες με την ιδιότητα $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$, είναι ακριβώς εκείνοι οι τετραγωνικοί πίνακες που διαθέτουν ένα πλήρες σύνολο από ορθοκανονικά ιδιοδιανύσματα και συνεπώς μπορούν να διαγωνοποιηθούν από ένα unitary πίνακα \mathbf{V} . Ειδικές περιπτώσεις κανονικών πινάκων είναι οι Hermitian πίνακες. Επομένως, το φασματικό θεώρημα της γραμμικής άλγεβρα δηλώνει ότι, κάθε Hermitian πίνακας \mathbf{A} δέχεται μια *αρμονική παραγοντοποίηση* ως

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i^H \quad (4)$$

όπου όλες οι ιδιοτιμές είναι πραγματικές. Εάν \mathbf{A} είναι Hermitian και θετικά ημι-ορισμένος, τότε οι ιδιοτιμές είναι μη-αρνητικές. Η αρμονική παραγοντοποίηση προφανώς εφαρμόζεται και σε κάθε *πραγματικό* συμμετρικό πίνακα, με την μόνη διαφορά ότι \mathbf{V} είναι ένας ορθογώνιος πίνακας.

Παραγοντοποίηση Συμμετρικών Θετικά-ορισμένων Πινάκων

Ένας τετραγωνικός πίνακας \mathbf{A} είναι Hermitian και θετικά ημι-ορισμένος εάν και μόνο εάν μπορεί να παραγοντοποιηθεί ως

$$\mathbf{A} = \mathbf{R}^H \mathbf{R} \quad (5)$$

όπου \mathbf{R} είναι οποιοσδήποτε πίνακας. Επιπλέον, ο \mathbf{A} είναι θετικά ορισμένος εάν και μόνο εάν ο \mathbf{R} έχει ανεξάρτητες στήλες. Τρεις επιλογές για τον \mathbf{R} είναι:

(1) Από την διάσπαση Cholesky του \mathbf{A} , μπορούμε να επιλέξουμε τον \mathbf{R} να είναι ο άνω τριγωνικός πίνακας $\sqrt{\mathbf{D}}\mathbf{L}^H$.

(2) Μια διαφορετική επιλογή προκύπτει από την αρμονική διάσπαση του \mathbf{A} , θέτοντας $\mathbf{R} = \sqrt{\mathbf{\Lambda}}\mathbf{V}^H$.

(3) Μια άλλη παραγοντοποίηση βασισμένη στην αρμονική παραγοντοποίηση είναι:

$$\mathbf{A} = \mathbf{R}^2, \quad \mathbf{R} = \mathbf{V}\sqrt{\mathbf{\Lambda}}\mathbf{V}^H \quad (6)$$

Η παραπάνω επιλογή για τον \mathbf{R} καλείται η Hermitian θετικά ημι-ορισμένη τετραγωνική ρίζα του \mathbf{A} .

Singular Value Decomposition (SVD) **&** **Ελαχιστα Τετραγωνα (Least Squares)**

Singular Value Decomposition (SVD): Κάθε (πραγματικός ή μιγαδικός) $m \times n$ πίνακας \mathbf{A} μπορεί να παραγοντοποιηθεί ως

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^H \quad (7)$$

όπου ο $m \times m$ πίνακας \mathbf{U} είναι unitary και οι στήλες του $\mathbf{u}_1, \dots, \mathbf{u}_m$ είναι τα ιδιοδιανύσματα του $\mathbf{A}\mathbf{A}^H$, ο $n \times n$ πίνακας \mathbf{V} είναι unitary και οι στήλες του $\mathbf{v}_1, \dots, \mathbf{v}_n$ είναι τα ιδιοδιανύσματα του $\mathbf{A}^H\mathbf{A}$, και ο $m \times n$ πίνακας $\mathbf{\Sigma}$ είναι πραγματικός διαγώνιος πίνακας του οποίου τα μονα μη-μηδενικά στοιχεία είναι οι r διαγώνιοι όροι του $\sigma_1, \sigma_2, \dots, \sigma_r > 0$, που καλούνται *ιδιόμορφες τιμές* (*singular values*), με

$$r = \text{rank}(\mathbf{A}) \leq \min(m, n). \quad (8)$$

Οι ιδιόμορφες τιμές είναι οι τετραγωνικές ρίζες των μη-μηδενικών ιδιοτιμών σ_i^2 και των δύο πινάκων $\mathbf{A}\mathbf{A}^H$ και $\mathbf{A}^H\mathbf{A}$. Επομένως, η SVD του \mathbf{A} σχετίζεται με την φασματική παραγοντοποίηση του Hermitian πίνακα $\mathbf{A}\mathbf{A}^H$ ως εξής:

$$\begin{aligned} \mathbf{A}\mathbf{A}^H &= \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^H = \sum_{i=1}^r \sigma_i^2 \mathbf{u}_i \mathbf{u}_i^H \\ \mathbf{A}^H\mathbf{A} &= \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{V}^H = \sum_{i=1}^r \sigma_i^2 \mathbf{v}_i \mathbf{v}_i^H \end{aligned} \quad (9)$$

Εάν ο \mathbf{A} είναι πραγματικός, η μόνη διαφορά στην SVD διάσπαση (συγκρινόμενη με την μιγαδική περίπτωση) είναι ότι οι \mathbf{U} και \mathbf{V} είναι ορθογώνιοι πίνακες. Εάν ο \mathbf{A} είναι Hermitian και θετικά ημι-ορισμένος, η SVD διάσπαση του θα είναι πανομοιότυπη με την φασματική διάσπαση $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$. Εάν ο \mathbf{A} είναι μη-ορισμένος, τότε κάθε αρνητική ιδιοτιμή στο \mathbf{A} γίνεται θετική στο $\mathbf{\Sigma}$.

Οι στήλες του πίνακα \mathbf{U} και \mathbf{V} παρέχουν ορθοκανονικές βάσεις για όλους τους τέσσερις βασικούς υποχώρους του \mathbf{A}
(Υποθέτουμε πραγματικούς πίνακες):

$$\begin{aligned}\mathcal{R}(\mathbf{A}) &= \text{column space,} & \dim &= r, & \text{eectors} &= \{\mathbf{u}_1, \dots, \mathbf{u}_r\} \\ \mathcal{R}(\mathbf{A}^T) &= \text{row space,} & \dim &= r, & \text{eectors} &= \{\mathbf{v}_1, \dots, \mathbf{v}_r\} \\ \mathcal{N}(\mathbf{A}) &= \text{null space,} & \dim &= n - r, & \text{eectors} &= \{\mathbf{v}_{r+1}, \dots, \mathbf{v}_n\} \\ \mathcal{N}(\mathbf{A}^T) &= \text{left null space,} & \dim &= m - r, & \text{eectors} &= \{\mathbf{u}_{r+1}, \dots, \mathbf{u}_m\}\end{aligned}\tag{10}$$

Αυτές οι βάσεις παρουσιάζουν ενδιάμεσες σχέσεις αφού

$$\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}\tag{11}$$

Επομένως, $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$ για $i = 1, \dots, n$. Για $i > r$, θέτουμε $\sigma_i = 0$.

Εφαρμογές της SVD:

(1) Αποδοτική Τάξη: Κρατούμε μόνο τις ιδιόμορφες τιμές που βρίσκονται πάνω από ένα καθορισμένο κατώφλι το οποίο καθορίζει την αριθμητική ακρίβεια.

(2) Συμπαγείς αναπαραστάσεις Εικόνων/Σημάτων: Χρήση μόνο λίγων μεγάλων ιδιόμορφων τιμών για να αναπαραστήσουμε προσεγγιστικά τον \mathbf{A} κάνοντας χρήση μιας εκδοχής της (7) που έχει υποστεί περικοπές.

(3) Πολική παραγοντοποίηση: Παραγοντοποίηση ενός πραγματικού τετραγωνικού πίνακα \mathbf{A} ως $\mathbf{Q}\mathbf{S}$ όπου ο \mathbf{Q} είναι ορθογώνιος και ο \mathbf{S} είναι συμμετρικός και θετικά ημι-ορισμένος. (Εάν ο \mathbf{A} είναι αντιστρέψιμος, ο \mathbf{S} είναι θετικά ορισμένος.)

$$\mathbf{A} = \mathbf{Q}\mathbf{S}, \quad \mathbf{Q} = \mathbf{U}\mathbf{V}^T, \quad \mathbf{S} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^T \quad (12)$$

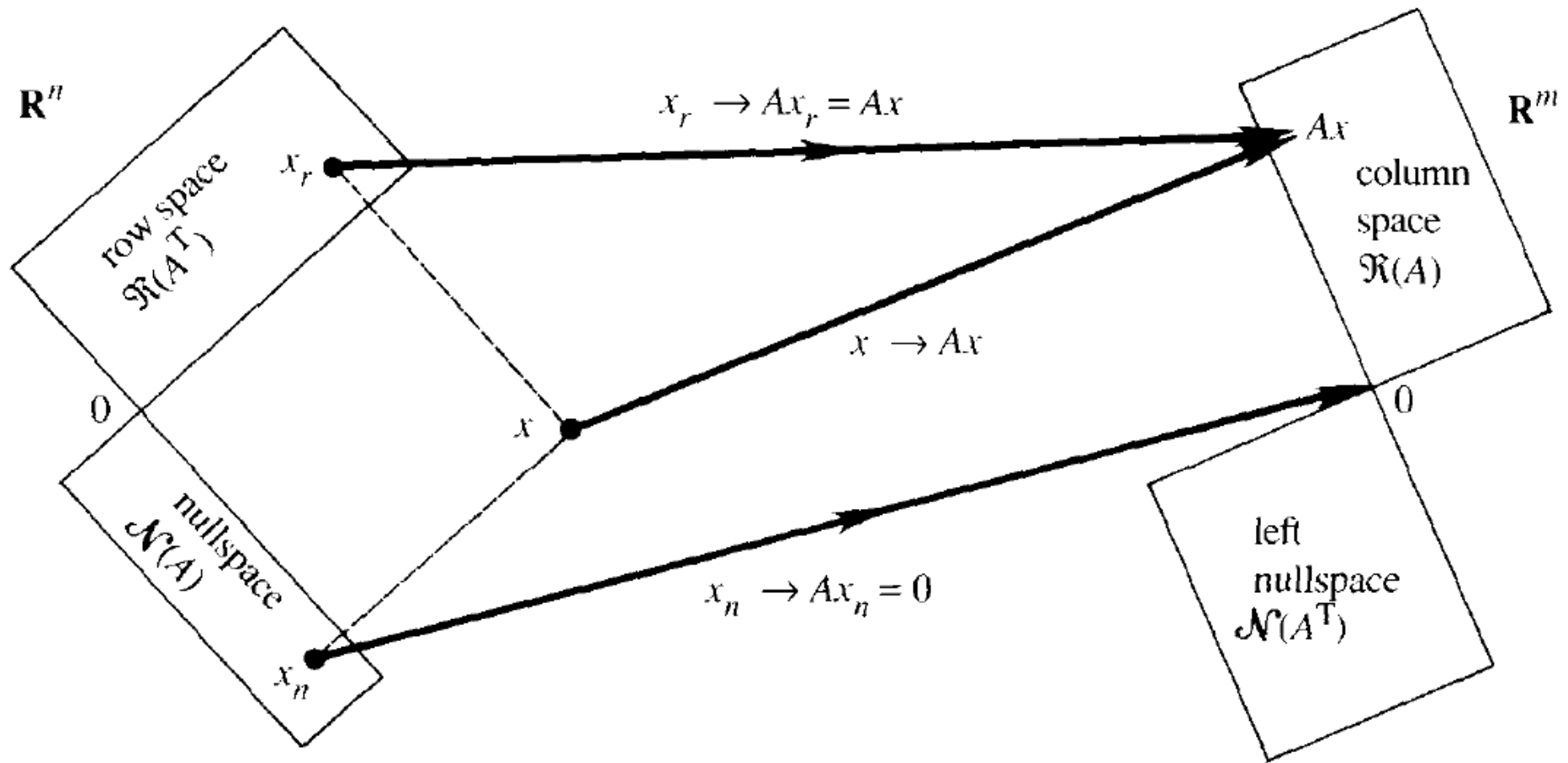
Αυτή η διάσπαση χρησιμοποιείται σε εφαρμογές ρομποτικής όπου ο \mathbf{Q} αναπαριστά περιστροφή ή ανάκλαση και ο \mathbf{S} κλιμάκωση (επέκταση ή συμπίεση).

(4) Ελάχιστα Τετράγωνα: Η **λύση ελαχίστων τετραγώνων** για το σύνολο των γραμμικών εξισώσεων $\mathbf{A}\mathbf{x} = \mathbf{b}$ δίνεται από

$$\mathbf{x}^+ = \mathbf{A}^+\mathbf{b}, \quad \mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^H \quad (13)$$

όπου \mathbf{A}^+ καλείται ο **ψευδο-αντίστροφος** του \mathbf{A} και ο $\mathbf{\Sigma}^+$ είναι ένας διαγώνιος πίνακας με $1/\sigma_1, \dots, 1/\sigma_r$ ως οι μόνοι μη-μηδενικοί διαγώνιοι του όροι.

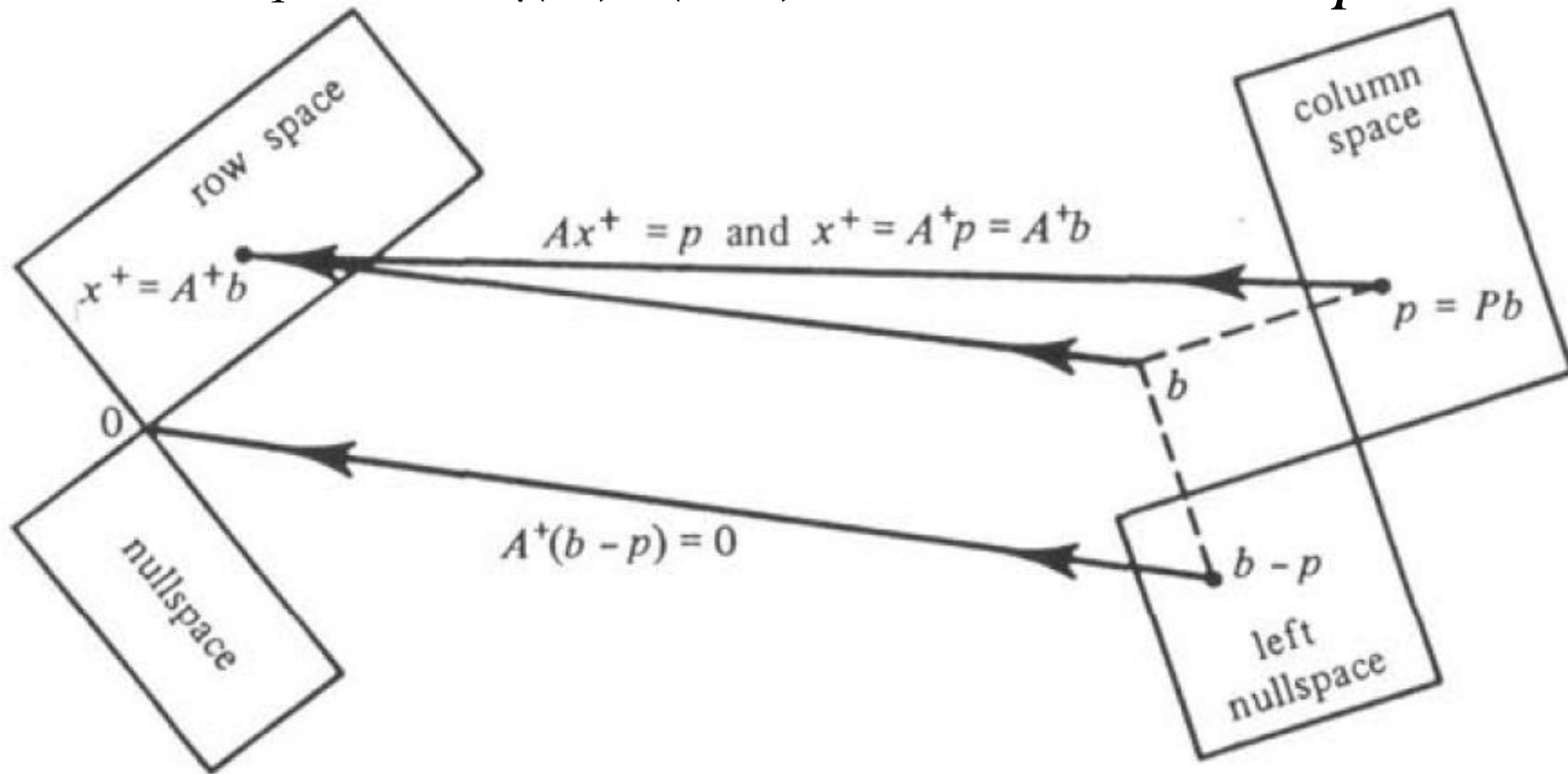
Δράση ενός πίνακα $m \times n$ A



Δραση του Ψευδο-αντιστροφου ενος πινακα A

Least Squares λυση(εις): $(A^T A)\hat{x} = A^T b$

$$A\hat{x} = p$$



$$\hat{x} = \hat{x}_r + \hat{x}_n, \quad A\hat{x}_r = p, \quad A\hat{x}_n = 0, \quad x^+ = \hat{x}_r$$

Προσεγγίση Πίνακα Δεδομένων με SVD

Έστω \mathbf{A} ένας πίνακας που παραγοντοποιούμε με τη χρήση της SVD και ταξινομούμε τις ιδιόμορφες τιμές ως $\sigma_1 \geq \dots \geq \sigma_r$. Εάν προσεγγίσουμε τον πίνακα διατηρώντας τις $p < r$ μεγαλύτερες ιδιόμορφες τιμές

$$\hat{\mathbf{A}} = \sum_{k=1}^p \sigma_k \mathbf{u}_k \mathbf{v}_k^H$$

η προσέγγιση αυτή μας οδηγεί στο μικρότερο τετραγωνικό σφάλμα

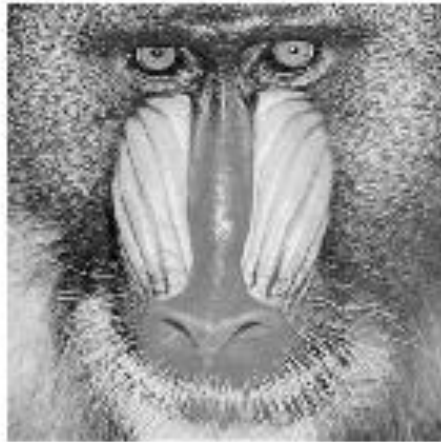
$$J_{svd} = \sum_{i=1}^m \sum_{j=1}^n |a[i, j] - \hat{a}[i, j]|^2 = (\|\mathbf{A} - \hat{\mathbf{A}}\|_{Frobenious})^2 = \sum_{k=p+1}^r \sigma_k^2$$

μεταξύ όλων των $m \times n$ πινάκων τάξης p . Συνεπώς, διατηρώντας τις p μεγαλύτερες ιδιόμορφες τιμές η SVD μας οδηγεί στην καλύτερη τάξης- p προσέγγιση πίνακα που ελαχιστοποιεί την Frobenius νόρμα του σφάλματος. Το ελάχιστο σφάλμα είναι

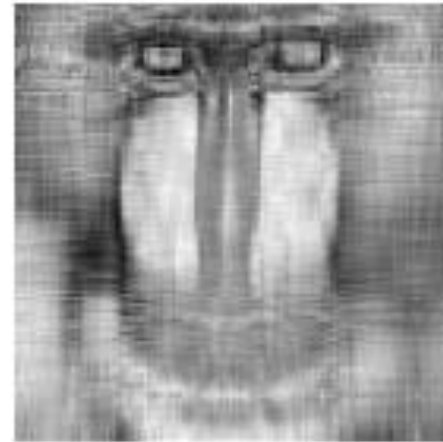
$$J_{svd} = \sum_{k=p+1}^r \sigma_k^2$$

Παραδειγμα SVD: Γκριζα Εικονα & Προσεγγιστικη Ανακατασκευη

Original Image



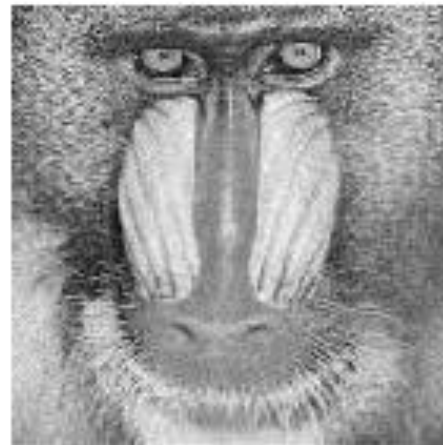
$k=10$



$k=20$



$k=50$



Παραδειγμα SVD: Εγχρωμη Εικόνα & Προσεγγιστική Ανακατασκευή

Original Image



$k=10$



$k=20$



$k=50$



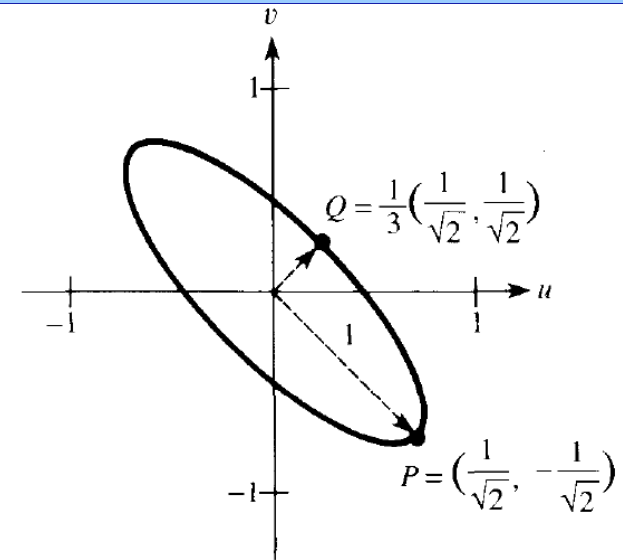
Θετικά-ορισμένοι Πινακες (Positive-definite Matrices)

- Γεωμετρία
- Rayleigh Quotient

Ελλιψοειδη και Θετικά Ορισμενοι Πινακες

$$A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

$$x^T A x = 5u^2 + 8uv + 5v^2 = 1$$



The ellipse $5u^2 + 8uv + 5v^2 = 1$ and its principal axes.

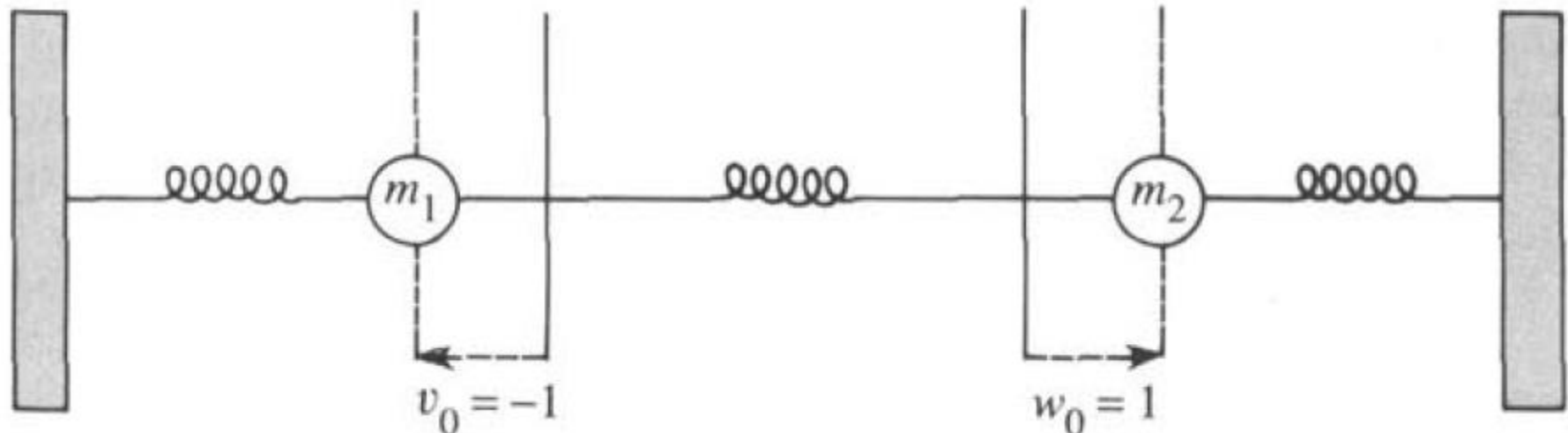
$$5u^2 + 8uv + v^2 = \left(\frac{u}{\sqrt{2}} - \frac{v}{\sqrt{2}} \right)^2 + 9 \left(\frac{u}{\sqrt{2}} + \frac{v}{\sqrt{2}} \right)^2 = 1.$$

6D Suppose A is positive definite: $A = Q\Lambda Q^T$ with $\lambda_i > 0$. Then the rotation $y = Q^T x$ simplifies $x^T A x = 1$ to

$$x^T Q \Lambda Q^T x = 1 \quad \text{or} \quad y^T \Lambda y = 1 \quad \text{or} \quad \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = 1.$$

This is the equation of an ellipsoid. Its axes have lengths $1/\sqrt{\lambda_1}, \dots, 1/\sqrt{\lambda_n}$ from the center, and in the original x -space they point along the eigenvectors.

Προβλήματα Γενικευμένων Ιδιοτιμών - 1



$$m_1 \frac{d^2 v}{dt^2} = -2v + w$$

$$m_2 \frac{d^2 w}{dt^2} = v - 2w$$

or

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \frac{d^2 u}{dt^2} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} u.$$

look for exponential solutions $e^{i\omega t}x$:

$$Mu'' = Au \quad \text{becomes} \quad M(i\omega)^2 e^{i\omega t}x = Ae^{i\omega t}x.$$

$$Ax = \lambda Mx, \quad \text{or} \quad \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} x = \lambda \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} x$$

Προβλήματα Γενικευμένων Ιδιοτιμών - 2

The underlying theory is easier to explain if M is split into $R^T R$. (M is assumed positive definite, as in the example.) Then the substitution $y = Rx$ changes

$$Ax = \lambda Mx = \lambda R^T R x \quad \text{into} \quad AR^{-1}y = \lambda R^T y.$$

Writing C for R^{-1} , and multiplying through by $(R^T)^{-1} = C^T$, this becomes a standard eigenvalue problem for the *single* matrix $C^T A C$:

$$C^T A C y = \lambda y. \quad (5)$$

The eigenvalues λ_j are the same as for the original $Ax = \lambda Mx$, and the eigenvectors are related by $y_j = Rx_j$.† The properties of the symmetric matrix $C^T A C$ lead directly to the corresponding properties of $Ax = \lambda Mx$:

1. The eigenvalues are real.
2. They have the same signs as the ordinary eigenvalues of A , by the law of inertia.
3. The eigenvectors y_j of $C^T A C$ can be chosen orthonormal, so the eigenvectors x_j are “ M -orthonormal”:

$$x_i^T M x_j = x_i^T R^T R x_j = y_i^T y_j = 1 \text{ if } i = j, 0 \text{ if } i \neq j. \quad (6)$$

4. Similarly $x_i^T A x_j = \lambda_j x_i^T M x_j$ is either λ_j or zero. The matrices A and M are being *simultaneously diagonalized*. If S has the x_j in its columns, then $S^T A S = \Lambda$ and $S^T M S = I$. Notice that it is a *congruence* transformation, with S^T on the left, and not a similarity transformation with S^{-1} .

$$x^T A x = z^T S^T A S z = \lambda_1 z_1^2 + \cdots + \lambda_n z_n^2 = 1$$

$$x^T M x = z^T S^T M S z = z_1^2 + \cdots + z_n^2 = 1.$$

Rayleigh Quotient

$$R(x) = \frac{x^T A x}{x^T x}.$$

61 *Rayleigh's principle:* The quotient $R(x)$ is minimized by the first eigenvector $x = x_1$, and its minimum value is the smallest eigenvalue λ_1 :

$$R(x_1) = \frac{x_1^T A x_1}{x_1^T x_1} = \frac{x_1^T \lambda_1 x_1}{x_1^T x_1} = \lambda_1.$$

Rayleigh Quotient: Απόδειξη

Geometrically, imagine that we fix the numerator at 1, and make the denominator as large as possible. The numerator $x^T A x = 1$ defines an ellipsoid, at least if A is positive definite. The denominator is $x^T x = \|x\|^2$, so we are looking for the point on the ellipsoid farthest from the origin—the vector x of greatest length. From our earlier description of the ellipsoid, its major axis points along the first eigenvector.

Algebraically, this is easy to see (without any requirement of positive definiteness). We diagonalize A by an orthogonal matrix: $Q^T A Q = \Lambda$. With $x = Qy$,

$$R(x) = \frac{(Qy)^T A (Qy)}{(Qy)^T (Qy)} = \frac{y^T \Lambda y}{y^T y} = \frac{\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2}{y_1^2 + \cdots + y_n^2}. \quad (3)$$

The minimum of R certainly occurs at the point where $y_1 = 1$ and $y_2 = \cdots = y_n = 0$. At that point the ratio equals λ_1 , and at any other point the ratio is larger:

$$\lambda_1(y_1^2 + y_2^2 + \cdots + y_n^2) \leq (\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2).$$

The Rayleigh quotient is *never below* λ_1 and *never above* λ_n . Its minimum is at the eigenvector x_1 and its maximum is at x_n . The intermediate eigenvectors are *saddle points*.

Γενικευμένο Rayleigh Quotient

Remark 2 For the generalized eigenvalue problem $Ax = \lambda Mx$, all the same principles hold if the denominator in the Rayleigh quotient is changed from $x^T x$ to $x^T Mx$. Substituting $x = Sz$, with the eigenvectors of $M^{-1}A$ in the columns of S , $R(x)$ simplifies to

$$R(x) = \frac{x^T A x}{x^T M x} = \frac{\lambda_1 z_1^2 + \cdots + \lambda_n z_n^2}{z_1^2 + \cdots + z_n^2}. \quad (9)$$

This was the point of simultaneously diagonalizing A and M in Section 6.3. Both quadratics become sums of perfect squares, and $\lambda_1 = \min R(x)$ as well as $\lambda_n = \max R(x)$. Even for *unequal* masses in an oscillating system, holding one mass at equilibrium will raise the lowest frequency and lower the highest frequency.

Βιβλιογραφία – Αναφορές

- G. Strang, *Linear Algebra and Its Applications*, Thomson Learning, 1988.
- M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, Wiley, 1996.