



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ

ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΣΟΦΙΑΣ ΜΑΡΑΓΚΟΥ

Παραγοντοποιήσεις Πινάκων και Εφαρμογές

Επιβλέπων : Φελλούρης Γ. Ανάργυρος

Αθήνα, Ιούλιος 2017



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ

ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΣΟΦΙΑΣ ΜΑΡΑΓΚΟΥ

Παραγοντοποιήσεις Πινάκων και Εφαρμογές

Επιτροπή : Ανάργυρος Φελλούρης (επιβλέπων)

Παναγιώτης Ψαρράκος

Πέτρος Στεφανέας

(Υπογραφή)

.....

Σοφία Μαραγκού

Διπλωματούχος Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών Ε.Μ.Π.

Παραγοντοποιήσεις Πινάκων και Εφαρμογές

Μαραγκού Σοφία

Αθήνα, Ιούλιος 2017

Περιεχόμενα

Ευχαριστίες	5
Εισαγωγή	6
Περίληψη	7
Abstract	8
Μέρος Α	9
(A1) Οι Στόχοι της Ανάλυσης Τυχαίου Πίνακα	9
(A2) Βασικά στοιχεία θεωρίας της Γραμμικής Άλγεβρας	10
Μέρος Β	21
1. Η LU – παραγοντοποίηση πίνακα	21
1.1 Ο αλγόριθμος της LU παραγοντοποίησης	21
1.2 Εφαρμογές και σχόλια	24
2. Η παραγοντοποίηση του Cholesky	25
2.1 Ο αλγόριθμος της παραγοντοποίησης Cholesky	25
2.2 Εφαρμογές και σχόλια	28
3. Η QR - παραγοντοποίηση πίνακα	31
3.1 Η μέθοδος ορθοκανονικοποίησης Gram-Schmidt	31
3.2 Ο αλγόριθμος της QR παραγοντοποίησης	34
3.3 Εφαρμογές και σχόλια	37
4. Η SVD - παραγοντοποίηση πίνακα	38
4.1 Ανάλυση ιδιζουσών τιμών	38
4.2 Ευστάθεια Τετραγωνικών Γραμμικών Συστημάτων	43
4.3 Ψευδοαντίστροφος Πίνακα	44
4.4 Εφαρμογές και σχόλια	50
Βιβλιογραφία	54

Ευχαριστίες

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω όλους εκείνους που συνέβαλαν, ο καθένας με τον τρόπο του, ώστε να γράφω αυτή την παράγραφο, αυτή τη στιγμή.

Τον επιβλέποντα καθηγητή μου, Ανάργυρο Φελλούρη, για την άψογη συνεργασία καθ' όλη τη διάρκεια της εκπόνησης της διπλωματικής μου εργασίας.

Τον πατέρα μου και κυρίως τη μητέρα μου που την έχασα κατά τη διάρκεια του τελευταίου χρόνου φοίτησής μου, αλλά αποτέλεσε για μένα πηγή δύναμης και αφοσίωσης στην επίτευξη των στόχων μου όλα αυτά τα χρόνια.

Την αδερφή μου και τους φίλους μου που πίστεψαν σε μένα και ήταν δίπλα μου στις δύσκολες στιγμές.

Εισαγωγή

Η παρούσα διπλωματική εργασία χωρίζεται σε δύο μέρη. Στο πρώτο (Α) μέρος εξηγούμε γιατί μας ενδιαφέρει η λεγόμενη ανάλυση σε παράγοντες ειδικούς πίνακες (decomposition) ή και απλά παραγοντοποίηση (factorization) ενός τυχαίου, όχι απαραίτητα τετραγωνικού πίνακα. Επιπλέον παρουσιάζονται κάποια βασικά στοιχεία θεωρίας της Γραμμικής Άλγεβρας. Στο δεύτερο (Β) μέρος παρουσιάζονται σε διαφορετικά κεφάλαια οι τέσσερις βασικές μέθοδοι παραγοντοποίησης πινάκων LU, Cholesky, QR και SVD αντίστοιχα, μαζί με παραδείγματα και τις εφαρμογές τους.

Περίληψη

Στόχος της παρούσας διπλωματικής εργασίας είναι η μελέτη των βασικών μεθόδων παραγοντοποίησης ενός τυχαίου $m \times n$ πίνακα και κατ' επέκταση οι σημαντικές εφαρμογές που προκύπτουν από την ανάλυση του πίνακα αυτού σε ειδικούς παράγοντες.

Οι βασικές μέθοδοι παραγοντοποίησης πίνακα είναι η μέθοδος LU, η μέθοδος του Cholesky η οποία αποτελεί ειδική περίπτωση της LU, η μέθοδος QR και τέλος η ανάλυση ιδιζουσών τιμών ή αλλιώς παραγοντοποίηση SVD. Οι εφαρμογές των μεθόδων αυτών είναι ποικίλες, με σημαντικότερες την επίλυση γραμμικών συστημάτων, τον υπολογισμό του αντίστροφου και ψευδοαντίστροφου πίνακα, την επίλυση του προβλήματος ελαχίστων τετραγώνων και τον υπολογισμό της τάξης πίνακα. Ακόμη σημαντικές είναι οι εφαρμογές στο χώρο της Στατιστικής, της Πληροφορικής και άλλων επιστημονικών κλάδων.

Στη μέθοδο LU παραγοντοποιούμε τον πίνακα A σε γινόμενο δύο πινάκων της μορφής $A = LU$, όπου ο L είναι ένας $m \times m$ κάτω τριγωνικός πίνακας με τα στοιχεία της κύριας διαγωνίου ίσα με 1, ενώ ο U είναι κλιμακωτός $m \times n$ πίνακας.

Η μέθοδος του Cholesky αποτελεί ειδική περίπτωση της LU παραγοντοποίησης, καθώς προϋποθέτει ο πίνακας A να είναι συμμετρικός και θετικά ορισμένος (όταν $K = \mathbb{R}$, αντίστοιχα ερμιτιανός και θετικά ορισμένος αν $K = \mathbb{C}$). Έτσι ο πίνακας A αναλύεται σε γινόμενο δύο πινάκων της μορφής $A = LL^T$ (αντίστοιχα $A = LL^*$), όπου ο πίνακας L είναι τριγωνικός κάτω και με θετικά στοιχεία στην κύρια διαγώνιο.

Στη μέθοδο QR ο πίνακας A είναι ένας $m \times n$ πίνακας με στήλες γραμμικώς ανεξάρτητα διανύσματα και μπορεί να γραφεί ως γινόμενο πινάκων της μορφής $A = QR$, όπου Q ο $m \times n$ πίνακας με στήλες ορθοκανονικά διανύσματα και R είναι ένας $n \times n$ αντιστρέψιμος τριγωνικός πάνω πίνακας.

Τέλος, στην ανάλυση ιδιζουσών τιμών (SVD) ο πίνακας $A \in M_{m \times n}(\mathbb{R})$ βαθμού $\rho(A) = r$ γράφεται ως γινόμενο $A = U\Sigma V^T$, (αντίστοιχα $A = U\Sigma V^*$ αν $A \in M_{m \times n}(\mathbb{C})$), όπου U είναι ένας $m \times m$ ορθογώνιος πίνακας, V είναι ένας $n \times n$ ορθογώνιος πίνακας και $\Sigma = (\sigma_{ij})$ είναι ένας $m \times n$ πίνακας με $\sigma_{ii} = \sigma_i$, $i = 1, 2, \dots, r$, $\sigma_{ii} = 0$, για $i = r+1, \dots, m$ και $\sigma_{ij} = 0$, για $i \neq j$, όπου σ_i οι ιδιάζουσες τιμές του πίνακα.

Abstract

The study of the basic factorization methods of a random $m \times n$ matrix and consequently the notable applications that result from the decomposition of this matrix constitute the purpose of the present thesis.

LU decomposition, Cholesky factorization (a subcategory of LU decomposition), QR method and finally Singular Value Decomposition (SVD) are the basic methods of matrix factorization. The aforementioned methods have various applications, the most significant of which are the linear systems solution, the calculation of inverse and pseudoinverse matrices, the solution of the least squares problem and finally finding the rank of the matrix. In addition, applications of decomposition methods in the fields of Statistics, Informatics, as well as in other scientific fields are of paramount importance.

LU decomposition of a matrix A factors as a product ($A=LU$) of a lower $m \times m$ triangular matrix L , whose elements of the main diagonal are equal to 1, and an echelon $m \times n$ matrix U .

Cholesky method is a special subcategory of LU decomposition, since it presupposes that, if $K = \mathbb{R}$, matrix A is symmetric and positive definite, and if $K = \mathbb{C}$, matrix A is Hermitian and positive definite. In this case, matrix A factors as a product $A = LL^T$ ($A = LL^*$ respectively), with matrix L being lower triangular and with positive elements in the main diagonal.

In method QR, A is a $m \times n$ matrix, with its columns being linearly independent vectors and factors as a product $A=QR$; Q is an $m \times n$ matrix with its columns being orthonormal vectors and R is an $n \times n$ invertible upper triangular matrix.

Finally, in Singular Value Decomposition (SVD), matrix $A \in M_{m \times n}(\mathbb{R})$, the rank of which is r , factors as a product $A = U\Sigma V^T$, (respectively $A = U\Sigma V^*$ if $A \in M_{m \times n}(\mathbb{C})$); U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix and $\Sigma = (\sigma_{ij})$ is an $m \times n$ matrix with $\sigma_{ii} = \sigma_i$, $i = 1, 2, \dots, r$, $\sigma_{ii} = 0$, for $i = r+1, \dots, m$ and $\sigma_{ij} = 0$, for $i \neq j$, with σ_i being the matrix singular values.

(A) ΜΕΡΟΣ

(A1) Οι Στόχοι της Ανάλυσης Τυχαίου Πίνακα

Σε πολλούς διαφορετικούς μαθηματικούς κλάδους (αλλά και στην κλασική φυσική) εμφανίζεται σε σημαντική έως μεγάλη έκταση μέσα στην ανάπτυξη μιας έννοιας της θεωρίας, όπως ο αριθμός, το πολυώνυμο, οι αλγεβρικές παραστάσεις κλπ, η επίτευξη μιας “κατάλληλης” παραγοντοποίησης. Έτσι, παραδείγματος χάριν, για έναν σύνθετο φυσικό αριθμό n οι αριθμοθεωρητικοί θέλουν - γρήγορα και με τον μικρότερο δυνατό κόπο αν γίνεται - να τον γράψουν στη μορφή $n = p_1^{k_1} p_2^{k_2} \dots p_m^{k_m}$, όπου p_1, p_2, \dots, p_m πρώτοι αριθμοί διαφορετικοί μεταξύ τους και k_1, k_2, \dots, k_m φυσικοί αριθμοί (η μηδέν). Βέβαια ο στόχος δεν είναι όπως βλέπουμε μια οποιαδήποτε παραγοντοποίηση! Κάτι ανάλογο συμβαίνει και με την άλγεβρα των πινάκων και ιδιαίτερα στις εφαρμογές τους: δεν θέλουμε μια οποιαδήποτε παραγοντοποίηση $A = BC$ ή BCD ενός πίνακα A , αλλά όπως πριν με την ανάλυση του n σε γινόμενο πρώτων έτσι κι εμείς θέλουμε να βρούμε τις ειδικές κατηγορίες των πινάκων αυτών, B, C και D που θα διευκολύνουν τις εφαρμογές των πινάκων.

Επιγραμματικά μερικές μόνο περιοχές εφαρμογής της παραγοντοποίησης των πινάκων είναι οι εξής:

- Επίλυση Γραμμικών Συστημάτων
- Υπολογισμός της τάξης πίνακα
- Δυνάμεις και αντιστρεψιμότητα πινάκων
- Πρόβλημα Ελαχίστων Τετραγώνων
- Στατιστικά μοντέλα
- Υπολογισμός ψευδοαντίστροφου

Οι εφαρμογές αυτές καθώς και άλλες θα εμφανιστούν στα κεφάλαια του (B) μέρους.

(A2) Βασικά στοιχεία θεωρίας της Γραμμικής Άλγεβρας

Πίνακες

Θεωρούμε τον πίνακα $A \in M_{m \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} , όπου έχουν οριστεί οι πράξεις της πρόσθεσης και του πολλαπλασιασμού. Το σύνολο $M_{m \times n}(K)$ εφοδιασμένο με τις πράξεις της πρόσθεσης και του πολλαπλασιασμού αριθμού με πίνακα αποκτά τη δομή διανυσματικού χώρου πάνω στο σώμα K .

Ο συμβολισμός για τον παραπάνω πίνακα είναι γνωστός

$$A := \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{bmatrix}$$

Αν $m = n$, τότε ο πίνακας λέγεται *τετραγωνικός*.

Ο τετραγωνικός πίνακας $A = (a_{ij}) \in M_{n \times n}$ λέγεται *άνω τριγωνικός*, αν όλα τα στοιχεία που βρίσκονται κάτω από την κύρια διαγώνιο είναι 0, δηλαδή είναι $a_{ij} = 0$, για $i > j$. Ομοίως, ο πίνακας $A = (a_{ij}) \in M_{n \times n}$ λέγεται *κάτω τριγωνικός*, αν όλα τα στοιχεία που βρίσκονται πάνω από την κύρια διαγώνιο είναι 0, δηλαδή είναι $a_{ij} = 0$, για $i < j$.

Το *γινόμενο δύο πινάκων* $A = (a_{ij}) \in M_{m \times n}(K)$, $B = (b_{ij}) \in M_{n \times p}(K)$ ορίζεται ως $C := AB \in M_{m \times p}(K)$ όπου $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$, $i = 1, \dots, m$, $j = 1, \dots, p$. Είναι γνωστό ότι η προσεταιριστική και η επιμεριστική ιδιότητα του πολλαπλασιασμού ισχύουν για πίνακες με την προϋπόθεση ότι οι πράξεις των πινάκων που σημειώνονται ορίζονται. Το ίδιο δεν ισχύει για την αντιμεταθετική ιδιότητα, ακόμη κι αν τα γινόμενα AB και BA που σημειώνονται ορίζονται. Είναι δηλαδή γενικά $AB \neq BA$.

Ο *μοναδιαίος πίνακας*

$$I_n = (\delta_{ij})_{n \times n} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \text{ όπου } \delta_{ij} = \begin{cases} 1, & \text{αν } i = j \\ 0, & \text{αν } i \neq j \end{cases}$$

Ορισμός: (Αντίστροφου πίνακα)

Αν ο πίνακας A είναι τετραγωνικός τύπου $n \times n$ και υπάρχει πίνακας X τέτοιος ώστε

$$AX = XA = I_n ,$$

Τότε λέμε ότι ο πίνακας A είναι *αντιστρέψιμος* και ο πίνακας X είναι ο *αντίστροφος* πίνακας του A . Γράφουμε τότε $X = A^{-1}$.

Αν οι $n \times n$ πίνακες A, B είναι αντιστρέψιμοι ισχύουν τα εξής:

$$\text{i) } (A^{-1})^{-1} = A \qquad \text{ii) } (AB)^{-1} = B^{-1}A^{-1} .$$

Ορισμός: (Δυνάμεων πίνακα)

Για κάθε τετραγωνικό $n \times n$ πίνακα A ορίζουμε τις δυνάμεις του ως εξής:

$$A^0 = I_n , A^1 = A \text{ και } A^n = A^{n-1}A , \text{ για κάθε } n \in \mathbb{N}^* \text{ με } n \geq 2$$

Ορισμός: (Ανάστροφου πίνακα)

Αν $A = (\alpha_{ij}) \in M_{m \times n}(K)$, τότε ο πίνακας που προκύπτει από τον A με εναλλαγή μεταξύ γραμμών και στηλών του λέγεται *ανάστροφος* πίνακας του A και συμβολίζεται A^T , έχουμε δηλαδή

$$A^T = (\alpha_{ji}) \in M_{n \times m} .$$

Ισχύουν τα εξής:

$$\text{i) } (A^T)^T = A , \text{ για κάθε } A \in M_{m \times n} \quad \text{ii) } (AB)^T = B^T A^T , \text{ για κάθε } A \in M_{m \times n} , B \in M_{n \times p}$$
$$\text{iii) } \text{όταν υπάρχει ο } A^{-1}, \text{ ισχύει: } (A^T)^{-1} = (A^{-1})^T$$

Ορισμός: (Συμμετρικού πίνακα)

Ο τετραγωνικός πίνακας $A = (\alpha_{ij}) \in M_n$ λέγεται:

- i) *Συμμετρικός*, αν $A^T = A$, δηλαδή αν $\alpha_{ij} = \alpha_{ji}$, για κάθε $i, j = 1, \dots, n$
- ii) *Αντισυμμετρικός*, αν $A^T = -A$, δηλαδή αν $\alpha_{ij} = -\alpha_{ji}$, για κάθε $i, j = 1, \dots, n$

Ορισμός: (Συζυγούς πίνακα)

Ο πίνακας \bar{A}^T λέγεται *συζυγής* πίνακας του $A \in M_n(\mathbb{C})$ και συμβολίζεται με A^* , ενώ όταν είναι $A \in M_n(\mathbb{R})$, τότε $A^* = A^T$.

Ισχύει $(AB)^* = B^* A^*$ και όταν υπάρχει ο A^{-1} , ισχύει: $(A^*)^{-1} = (A^{-1})^*$.

Ορισμός: (Ερμιτιανού πίνακα)

Ο πίνακας $A \in M_n(\mathbb{C})$ λέγεται *ερμιτιανός*, αν ικανοποιεί την ιδιότητα

$$A^* := \bar{A}^T = A$$

Ενώ ο πίνακας $A \in M_n(\mathbb{R})$ λέγεται *συμμετρικός* αν ισχύει

$$A^* := A^T = A$$

Ορισμός: (Ορθομοναδιαίου και Ορθογώνιου πίνακα)

Ο πίνακας $A \in M_n(\mathbb{C})$ λέγεται *ορθομοναδιαίος* (unitary), αν ικανοποιεί τις ισότητες $AA^* = A^*A = I$ ή ισοδύναμα

$$A^* = A^{-1}.$$

Ο πίνακας $A \in M_n(\mathbb{R})$ λέγεται *ορθογώνιος* αν ικανοποιεί τις ισότητες $AA^T = A^T A = I$ ή ισοδύναμα

$$A^T = A^{-1}.$$

Ορισμός: (Όμοιοι πίνακες)

Ο πίνακας $A \in M_n(K)$, $K = \mathbb{R}$ ή \mathbb{C} λέγεται *όμοιος* με τον πίνακα $B \in M_n(K)$, αν υπάρχει αντιστρέψιμος πίνακας P τέτοιος ώστε να ισχύει

$$B = P^{-1}AP.$$

Σημειώνουμε ότι, σε μια γραμμή ενός πίνακα $A = (a_{ij}) \in M_{m \times n}(K)$ η οποία δεν έχει όλα τα στοιχεία της 0 (μη μηδενική γραμμή), το ηγετικό (ή βασικό) στοιχείο της είναι το πρώτο, από αριστερά προς τα δεξιά, μη μηδενικό στοιχείο της.

Ορισμός: (Κλιμακωτού πίνακα)

Ο πίνακας $A = (a_{ij}) \in M_{m \times n}(K)$ λέγεται κλιμακωτός ή κλιμακωτής μορφής, όταν ισχύουν τα ακόλουθα:

- i) Οι μηδενικές γραμμές, αν υπάρχουν, βρίσκονται στη σειρά μετά τις μη μηδενικές γραμμές και
- ii) Αν $\gamma_1, \gamma_2, \dots, \gamma_k$, $k \leq m$ είναι οι μη μηδενικές γραμμές του πίνακα A , τότε το ηγετικό στοιχείο της γ_{i+1} – γραμμής βρίσκεται δεξιότερα από το ηγετικό στοιχείο της γ_i – γραμμής.

Παράδειγμα κλιμακωτού πίνακα: $A = \begin{bmatrix} 3 & 0 & 1 & -2 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Ορισμός: (Ανηγμένου κλιμακωτού πίνακα)

Ο πίνακας $A = (a_{ij}) \in M_{m \times n}(K)$ λέγεται ανηγμένος κλιμακωτός ή ανηγμένης κλιμακωτής μορφής, αν ισχύουν τα ακόλουθα:

- i) Είναι κλιμακωτός,
- ii) Το ηγετικό στοιχείο κάθε μη μηδενικής γραμμής είναι 1 και
- iii) Σε μία στήλη που περιέχει το ηγετικό στοιχείο κάποιας γραμμής, όλα τα στοιχεία της είναι 0.

Παράδειγμα ανηγμένου κλιμακωτού πίνακα: $A = \begin{bmatrix} 1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

Στοιχειώδεις πράξεις και πίνακες

Στο σύνολο των $m \times n$ πινάκων θεωρούμε απεικονίσεις της μορφής

$$\tau : M_{m \times n} \rightarrow M_{m \times n}, A \rightarrow \tau(A),$$

που λέγονται *στοιχειώδεις πράξεις γραμμών* ή *γραμμοπράξεις*.

Ο πίνακας $\tau(A)$ προκύπτει από τον πίνακα A με εφαρμογή μιας από τις ακόλουθες διαδικασίες:

- i) $\gamma_i \leftrightarrow \gamma_j$, εναλλαγή i – γραμμής με την j – γραμμή
- ii) $\gamma_i \leftrightarrow \lambda \gamma_i$, $\lambda \neq 0$, πολλαπλασιασμός της i – γραμμής επί $\lambda \neq 0$.
- iii) $\gamma_i \leftrightarrow \gamma_i + \lambda \gamma_j$, $\lambda \neq 0$, αντικατάσταση της της i – γραμμής από το άθροισμα αυτής και του λ – πλάσιου της j – γραμμής.

Αντίστοιχα ορίζονται και πράξεις στηλών ή στηλοπράξεις.

Σε κάθε μία από τις στοιχειώδεις πράξεις γραμμών (αντίστοιχα στηλών) αντιστοιχίζουμε ένα *στοιχειώδη* πίνακα. Έτσι, στη στοιχειώδη πράξη γραμμών τ που εφαρμόζεται σε $m \times n$ πίνακα, αντιστοιχίζουμε τον στοιχειώδη πίνακα $\tau(I_m)$ που προκύπτει από την εφαρμογή της τ στο μοναδιαίο πίνακα I_m . Για παράδειγμα στις στοιχειώδεις πράξεις

$$\gamma_1 \leftrightarrow \gamma_2, \gamma_3 \leftrightarrow \gamma_3 + 2\gamma_1$$

όταν εφαρμόζονται σε 3×4 πίνακα, αντιστοιχίζονται οι παρακάτω στοιχειώδεις πίνακες:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}.$$

Γραμμικά συστήματα

Το γραμμικό σύστημα με m εξισώσεις και n αγνώστους ή $m \times n$ – γραμμικό σύστημα γράφεται ως μια εξίσωση πινάκων

$$AX = B,$$

όπου ο πίνακας A ($m \times n$) είναι πίνακας συντελεστών, ο πίνακας X ($n \times 1$) είναι ο πίνακας των αγνώστων και ο πίνακας B ($m \times 1$) είναι ο πίνακας των σταθερών όρων.

Λύση του γραμμικού συστήματος λέγεται κάθε διατεταγμένη n – άδα $(\xi_1, \dots, \xi_n) \in \mathbb{K}^n$ που επαληθεύει και τις m εξισώσεις του συστήματος.

Το γραμμικό σύστημα λέγεται *συμβιβαστό*, αν έχει μία τουλάχιστον λύση, ενώ λέγεται *αδύνατο*, αν δεν έχει λύση.

Αν ο πίνακας B των σταθερών όρων είναι ο μηδενικός πίνακας, τότε το γραμμικό σύστημα γίνεται

$$AX = 0$$

και λέγεται *ομογενές*.

Θεώρημα:

Αν $A \in \mathbb{C}^{n \times n}$, τότε οι παρακάτω προτάσεις είναι ισοδύναμες:

- i) Υπάρχει ο A^{-1}
- ii) Τα διανύσματα – στήλες του A είναι γραμμικώς ανεξάρτητα
- iii) $\det(A) \neq 0$
- iv) Το γραμμικό σύστημα $AX = 0$ έχει τη μοναδική λύση $X = 0$
- v) Το γραμμικό σύστημα $AX = B$ έχει τη μοναδική λύση $X = A^{-1}B$

Ιδιοτιμές και Ιδιοδιανύσματα πίνακα

Είναι γνωστό ότι για κάθε $A \in M_{n \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} υπάρχουν $\lambda_i \in \mathbb{R}$ ή \mathbb{C} και $x^i \in \mathbb{R}^n \setminus \{0\}$ ή $\mathbb{C}^n \setminus \{0\}$, $i = 1, \dots, n$ τέτοια ώστε

$$Ax^i = \lambda_i x^i, \quad i = 1, \dots, n.$$

Οι αριθμοί λ_i καλούνται *ιδιοτιμές* και τα x^i (αντίστοιχα) *ιδιοδιανύσματα* του πίνακα A .

Θεώρημα:

Ο αριθμός $\lambda \in K$ είναι ιδιοτιμή του πίνακα $A \in M_{n \times n}(K)$, αν και μόνο αν, ικανοποιεί την εξίσωση (ως προς t)

$$\det(tI - A) = 0.$$

Τα ιδιοδιανύσματα προκύπτουν από την επίλυση του ομογενούς γραμμικού συστήματος $(\lambda_i I - A)x^i = 0$.

Το σύνολο των ιδιοτιμών του πίνακα A λέγεται *φάσμα* του A και συμβολίζεται με $\sigma(A)$.

Επιπλέον ονομάζουμε *φασματική ακτίνα* του πίνακα A και συμβολίζουμε με $\rho(A)$ τον αριθμό

$$\rho(A) := \max_{i=1}^n |\lambda_i|.$$

Θεώρημα:

Ένας συμμετρικός πίνακας A είναι *θετικά (αρνητικά) ορισμένος* αν και μόνο αν, όλες οι ιδιοτιμές του πίνακα A είναι θετικές (αρνητικές). Αντίστοιχα είναι *θετικά ημιορισμένος* αν και μόνο αν, όλες οι ιδιοτιμές του πίνακα A είναι μη αρνητικές.

Θεώρημα:

Ένας συμμετρικός πίνακας $A = (a_{ij})$ είναι θετικά ορισμένος αν και μόνο αν, $A_j > 0$, για κάθε $j = 1, 2, \dots, n$, όπου A_j είναι οι κύριες υποορίζουσες του πίνακα A που ορίζονται από

$$A_1 = a_{11}, A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \dots, A_n = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \det A.$$

Ένας συμμετρικός πίνακας $A = (a_{ij})$ είναι αρνητικά ορισμένος αν και μόνο αν, ισχύει $(-1)^j A_j > 0$, για κάθε $j = 1, 2, \dots, n$.

Συνεχίζοντας, θεωρούμε γνωστές μερικές από τις βασικές έννοιες της Γραμμικής Άλγεβρας. Ιδιαίτερα θεωρούνται γνωστά τα περί: “Διανυσματικών Χώρων” και “Διάστασης Διανυσματικού Χώρου”. Ειδικότερα θα μας απασχολήσουν οι Διανυσματικοί Χώροι \mathbb{R}^n και \mathbb{C}^n , για τα διανύσματα των οποίων θα χρησιμοποιηθούν οι παρακάτω συμβολισμοί:

$$x := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, x_i \in \mathbb{C} \text{ ή } \mathbb{R}, i = 1, 2, \dots, n$$

$x^T = [x_1 \ x_2 \ \dots \ x_n]$, ανάστροφο του x .

$x^* = [\bar{x}_1 \ \bar{x}_2 \ \dots \ \bar{x}_n]$, συζυγές ανάστροφο του x .

Ορισμός: (Νόρμας)

Έστω V διανυσματικός χώρος πάνω στο σώμα K . Μία απεικόνιση $\|\cdot\| : V \rightarrow \mathbb{R}$ ονομάζεται *νόρμα* αν για κάθε $x, y \in V$ και $\lambda \in K$, ικανοποιεί τα ακόλουθα:

- i. $\|x\| \geq 0$ για κάθε $x \in V$
- ii. $\|x\| = 0$ αν και μόνο αν $x = 0$
- iii. $\|\lambda x\| = |\lambda| \|x\|$
- iv. $\|x + y\| \leq \|x\| + \|y\|$ (τριγωνική ανισότητα)

Άμεση συνέπεια των δύο τελευταίων ιδιοτήτων και η $|\|x\| - \|y\|| \leq \|x - y\|$.

Ορισμός: (Εσωτερικού γινομένου)

Έστω V διανυσματικός χώρος πάνω στο σώμα $K = \mathbb{R}$ ή \mathbb{C} . Μία απεικόνιση $\langle \cdot | \cdot \rangle : V \times V \rightarrow K$ ονομάζεται *εσωτερικό γινόμενο* αν για κάθε $x, y, z \in V$ και $\lambda, \mu \in K$, ικανοποιεί τα ακόλουθα:

- i. $\langle x | x \rangle \geq 0$ για κάθε $x \in V$
- ii. $\langle x | x \rangle = 0$ αν και μόνο αν $x = 0$
- iii. $\langle \lambda x + \mu y | z \rangle = \lambda \langle x | z \rangle + \mu \langle y | z \rangle$
- iv. $\langle x | y \rangle = \overline{\langle y | x \rangle}$
Για $K = \mathbb{R}$, η iv) γίνεται $\langle x | y \rangle = \langle y | x \rangle$.

Αν είναι $K = \mathbb{R}$ και $\dim V < +\infty$, τότε ο χώρος εσωτερικού V λέγεται *Ευκλείδειος*, ενώ αν είναι $K = \mathbb{C}$ και $\dim V < +\infty$, τότε ο χώρος εσωτερικού V λέγεται *Ορθομοναδιαίος (unitary)*.

Από τον ορισμό του εσωτερικού γινομένου, εύκολα μπορεί κανείς να επιβεβαιώσει τις παρακάτω ιδιότητες:

1. $\langle x | \alpha y \rangle = \overline{\langle \alpha y | x \rangle} = \bar{\alpha} \overline{\langle y | x \rangle} = \bar{\alpha} \langle x | y \rangle$
2. $\langle x | y + z \rangle = \overline{\langle y + z | x \rangle} = \overline{\langle y | x \rangle} + \overline{\langle z | x \rangle} = \langle x | y \rangle + \langle x | z \rangle$
3. $x = 0$ αν και μόνο αν $\langle x | y \rangle = 0$ για κάθε $y \in V$.

Αν $\|\cdot\|$ είναι μία νόρμα διανυσμάτων τέτοια ώστε να ισχύει $\|x\| = \sqrt{\langle x|x \rangle}$ για κάθε $x \in V$, για κάποιο εσωτερικό γινόμενο $\langle \cdot | \cdot \rangle$, τότε λέμε ότι η $\|\cdot\|$ επάγεται από το εσωτερικό γινόμενο $\langle \cdot | \cdot \rangle$.

Στην περίπτωση αυτή, για κάθε $x, y \in V$ ισχύει

$$\begin{aligned} \frac{1}{2} (\|x + y\|^2 + \|x - y\|^2) &= \frac{1}{2} (\langle x + y | x + y \rangle + \langle x - y | x - y \rangle) \\ &= \frac{1}{2} (2\langle x | x \rangle + 2\langle y | y \rangle). \end{aligned}$$

Δηλαδή, η νόρμα $\|\cdot\|$ ικανοποιεί τον κανόνα του παραλληλογράμμου :

$$\frac{1}{2} (\|x + y\|^2 + \|x - y\|^2) = \|x\|^2 + \|y\|^2, \text{ για κάθε } x, y \in V.$$

Μια σημαντική ιδιότητα του εσωτερικού γινομένου είναι η ανισότητα Cauchy – Schwarz.

Θεώρημα:

Έστω $\langle \cdot | \cdot \rangle$ ένα εσωτερικό γινόμενο στον διανυσματικό χώρο V και $\|\cdot\|$ η επαγόμενη νόρμα. Τότε για κάθε $x, y \in V$, ισχύει

$$|\langle x | y \rangle|^2 \leq \langle x | x \rangle \langle y | y \rangle,$$

ή ισοδύναμα,

$$|\langle x | y \rangle| \leq \|x\| \|y\|$$

Η ισότητα ισχύει αν και μόνο αν τα διανύσματα x και y είναι γραμμικώς εξαρτημένα, δηλαδή $x = \lambda y$.

Παραδείγματα χώρων με εσωτερικό γινόμενο:

- Ο Ευκλείδειος διανυσματικός χώρος \mathbb{R}^n με το κανονικό εσωτερικό γινόμενο ή Ευκλείδειο εσωτερικό γινόμενο

$$\langle x|y \rangle = \sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

- Ο Ορθομοναδιαίος διανυσματικός χώρος \mathbb{C}^n με το κανονικό εσωτερικό γινόμενο

$$\langle x|y \rangle = \sum_{i=1}^n x_i \bar{y}_i = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n$$

Τέσσερα χαρακτηριστικά παραδείγματα νορμών διανυσμάτων στο \mathbb{C}^n που χρησιμοποιούνται συχνά είναι τα ακόλουθα:

- Η l_p -νόρμα, για οποιονδήποτε πραγματικό αριθμό $p \geq 1$, ορίζεται ως

$$\|x\|_p = \|[x_1 \ x_2 \ \dots \ x_n]^T\|_p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{1/p}$$

- Η Ευκλείδεια νόρμα (ή l_2 -νόρμα) ορίζεται ως

$$\|x\|_2 = \|[x_1 \ x_2 \ \dots \ x_n]^T\|_2 = (|x_1|^2 + |x_2|^2 + \dots + |x_n|^2)^{1/2}$$

Αποτελεί μια ειδική περίπτωση της p -νόρμας (για $p = 2$), είναι ίσως η πιο γνωστή νόρμα διανυσμάτων και επάγεται από το κανονικό εσωτερικό γινόμενο.

- Η αθροιστική νόρμα (ή l_1 -νόρμα) ορίζεται ως

$$\|x\|_1 = \|[x_1 \ x_2 \ \dots \ x_n]^T\|_1 = |x_1| + |x_2| + \dots + |x_n|$$

- Η μέγιστη νόρμα (ή max-νόρμα ή ∞ -νόρμα) ορίζεται ως

$$\|x\|_\infty = \|[x_1 \ x_2 \ \dots \ x_n]^T\|_\infty = \max \{ |x_1|, |x_2|, \dots, |x_n| \}.$$

(B) ΜΕΡΟΣ

1. Η LU – παραγοντοποίηση πίνακα

1.1 Ο αλγόριθμος της LU παραγοντοποίησης

Στην παράγραφο αυτή παρουσιάζεται μία μέθοδος επίλυσης του γραμμικού συστήματος $AX = B$ (Σ) η οποία βασίζεται στην παραγοντοποίηση του πίνακα A των συντελεστών σε γινόμενο δύο πινάκων, ενός κάτω τριγωνικού και ενός κλιμακωτού πίνακα.

Ας υποθέσουμε ότι έχουμε να λύσουμε το γραμμικό σύστημα (Σ). Θεωρούμε πίνακα $A \in M_{m \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} και με στοιχειώδεις πράξεις γραμμών, χωρίς να κάνουμε εναλλαγή γραμμών, μετατρέπουμε τον πίνακα σε κλιμακωτό, οπότε έχουμε

$$P_5 \dots P_2 P_1 A = U, \quad (1)$$

όπου P_1, P_2, \dots, P_5 είναι οι αντίστοιχοι στοιχειώδεις πίνακες. Επειδή όλοι οι στοιχειώδεις πίνακες είναι αντιστρέψιμοι έχουμε

$$A = P_1^{-1} P_2^{-1} \dots P_5^{-1} U = LU, \quad (2)$$

όπου έχουμε θέσει $L = P_1^{-1} P_2^{-1} \dots P_5^{-1}$. Επιπλέον λαμβάνουμε

$$P_5 \dots P_2 P_1 L = I, \quad (3)$$

οπότε παρατηρούμε ότι οι ίδιες γραμμοπράξεις μετατρέπουν τον πίνακα L στον μοναδιαίο πίνακα.

Με αυτή τη διαδικασία παραγοντοποιούμε τον πίνακα A σε γινόμενο δύο πινάκων της μορφής

$$A = LU, \quad (4)$$

όπου ο L είναι $m \times m$ κάτω τριγωνικός με τα στοιχεία της κύριας διαγωνίου ίσα με 1, ενώ ο U είναι κλιμακωτός $m \times n$ πίνακας.

Έχουμε

$$AX = B \Leftrightarrow (LU)X = B \Leftrightarrow L(UX) = B,$$

οπότε αν θέσουμε $Y = UX$, τότε έχουμε να επιλύσουμε τα απλά γραμμικά συστήματα με πίνακα κλιμακωτής μορφής

$$LY = B \quad \text{και} \quad UX = Y.$$

Επιλύουμε το σύστημα $LY = B$ με αντικατάσταση προς τα εμπρός και στη συνέχεια επιλύουμε το σύστημα $UX = Y$ με αντικατάσταση προς τα πίσω.

Για παράδειγμα, ας υποθέσουμε πως έχουμε προς λύση το σύστημα

$$AX = \begin{bmatrix} 2 & 1 & -1 & 0 \\ 4 & 3 & 0 & -2 \\ 2 & -2 & -9 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = B$$

Έχουμε

$$\begin{bmatrix} 2 & 1 & -1 & 0 \\ 4 & 3 & 0 & -2 \\ 2 & -2 & -9 & 8 \end{bmatrix} \xrightarrow{\substack{\gamma_2 \rightarrow \gamma_2 - 2\gamma_1 \\ \gamma_3 \rightarrow \gamma_3 - \gamma_1}} \begin{bmatrix} 2 & 1 & -1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & -3 & -8 & 8 \end{bmatrix}.$$

Μέχρι τώρα χρησιμοποιήσαμε ως βασικό (pivot) στοιχείο το $a_{11} = 2$ (συμβολίζεται με παχύ μαύρο) και παρατηρούμε ότι οι ίδιες στοιχειώδεις πράξεις δίνουν

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & * & 1 \end{bmatrix} \xrightarrow{\substack{\gamma_2 \rightarrow \gamma_2 - 2\gamma_1 \\ \gamma_3 \rightarrow \gamma_3 - \gamma_1}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & * & 1 \end{bmatrix}.$$

Στη συνέχεια έχουμε

$$\begin{bmatrix} 2 & 1 & -1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & -3 & -8 & 8 \end{bmatrix} \xrightarrow{\gamma_3 \rightarrow \gamma_3 + 3\gamma_2} \begin{bmatrix} 2 & 1 & -1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & -2 & 2 \end{bmatrix} = U,$$

όπου χρησιμοποιήσαμε ως βασικό στοιχείο $a'_{22} = 1$. Παρατηρούμε ότι η ίδια στοιχειώδης γραμμοπράξη δίνει

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} \xrightarrow{\gamma_3 \rightarrow \gamma_3 + 3\gamma_2} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

οπότε πλέον είναι φανερό ότι ο κατάλληλος πίνακας L που με τις ίδιες γραμμοπράξεις μετατρέπεται στο μοναδιαίο είναι ο

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -3 & 1 \end{bmatrix}.$$

Μπορούμε πλέον να περιγράψουμε τη διαδικασία παραγοντοποίησης του πίνακα A ως εξής:

- i) Μετατρέπουμε τον πίνακα A με στοιχειώδεις πράξεις γραμμών, χωρίς εναλλαγές γραμμών, σε κλιμακωτό, οπότε προκύπτει ο πίνακας U.
- ii) Διαμορφώνουμε κάθε στήλη του A που έχει βασικό στοιχείο, κάθε φορά που εμφανίζεται, με μηδενικά πάνω από το βασικό στοιχείο και διαιρώντας τα υπόλοιπα στοιχεία της με το βασικό στοιχείο. Οι στήλες που διαμορφώνονται μαζί με μία τελευταία στήλη της μορφής $[0 \dots 0 1]^T$ αποτελούν τις στήλες του πίνακα L. (Με απλά λόγια ο πίνακας L αποτελείται από τους πολλαπλασιαστές που χρησιμοποιήθηκαν σε κάθε βήμα και αποτελούν τις στήλες του. Η τελευταία στήλη είναι της μορφής $[0 \dots 0 1]^T$ και τα στοιχεία της κύριας διαγωνίου του πίνακα L ισούνται με 1.)

Αφού έχουμε βρει την παραγοντοποίηση

$$A = \begin{bmatrix} 2 & 1 & -1 & 0 \\ 4 & 3 & 0 & -2 \\ 2 & -2 & -9 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & -2 & 2 \end{bmatrix} = LU$$

προχωράμε στην επίλυση του συστήματος.

Επιλύουμε πρώτα το σύστημα $LY = B$,

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -3 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \Leftrightarrow Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ -8 \end{bmatrix}$$

Στην συνέχεια επιλύουμε το σύστημα $UX = Y$

$$\begin{bmatrix} 2 & 1 & -1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ -8 \end{bmatrix} \Leftrightarrow X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 8,5 + 0,5\alpha \\ -11 \\ 4 + \alpha \\ \alpha \end{bmatrix}, \alpha \in \mathbb{R} .$$

1.2 Εφαρμογές και σχόλια

Ειδική περίπτωση των πινάκων $A \in M_{m \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} είναι οι τετραγωνικοί πίνακες $A \in M_{n \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} . Η παραγοντοποίηση LU ενός τετραγωνικού πίνακα A είναι δυνατή όταν δεν παρουσιάζονται μηδενικά οδηγά στοιχεία και συνεπώς δεν απαιτείται η εναλλαγή γραμμών. Ο αντιστρέψιμος (σε αυτήν την περίπτωση) πίνακας $A \in M_{n \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} γράφεται στη μορφή $A = LU$, όπου ο πίνακας L είναι κάτω τριγωνικός με όλα τα διαγώνια στοιχεία του ίσα με 1, και ο πίνακας U είναι άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία (γνωστά ως βασικά-ρινοτς στοιχεία). Μία προφανής χρήση της ανάλυσης LU του πίνακα A είναι η ευκολία στον υπολογισμό του A^{-1} .

Η παραγοντοποίηση ενός πίνακα A στη μορφή LU έχει μεγάλη πρακτική σημασία στην επίλυση γραμμικών συστημάτων. Η εξίσωση $AX = B \Leftrightarrow (LU)X = B$ επιλύεται εύκολα σε δύο βήματα. Αρχικά λύνουμε το σύστημα $LY = B$ και στη συνέχεια το σύστημα $UX = Y$. Τα δύο νέα συστήματα είναι τριγωνικά και συνεπώς εύκολο να λυθούν.

Επιπλέον το κέρδος της παραγοντοποίησης LU είναι σημαντικό όταν πρέπει να λύσουμε ταυτόχρονα περισσότερα από ένα γραμμικά συστήματα με τον ίδιο πίνακα συντελεστών A και διαφορετικό πίνακα σταθερών όρων. Είναι κατάλληλη για ηλεκτρονικούς υπολογιστές και αποτελεί βάση πολλών υπολογιστικών πακέτων.

2. Η παραγοντοποίηση του Cholesky

2.1 Ο αλγόριθμος της παραγοντοποίησης Cholesky

Η παραγοντοποίηση του Cholesky είναι ειδική περίπτωση της LU – παραγοντοποίησης πίνακα που αναλύσαμε στο προηγούμενο κεφάλαιο.

Ένας συμμετρικός ($A^T = A$) και θετικά ορισμένος πίνακας A μπορεί να αναλυθεί σε γινόμενο δύο πινάκων της μορφής $A = LL^T$, όπου ο πίνακας L είναι τριγωνικός κάτω και με θετικά στοιχεία στην κύρια διαγώνιο. Έτσι για την επίλυση του γραμμικού συστήματος $AX = B$, αρκεί πρώτα να επιλύσουμε ως προς Y το σύστημα $LY = B$ και στη συνέχεια ως προς X το σύστημα $L^T X = Y$.

Δηλαδή,

$$AX = B \Leftrightarrow (LL^T)X = B \Leftrightarrow L(L^T X) = B$$

οπότε αν θέσουμε $Y = L^T X$ επιλύουμε ως προς Y το σύστημα $LY = B$ και στη συνέχεια ως προς X το σύστημα $L^T X = Y$.

Αν $A = (a_{ij}) \in M_n(\mathbb{R})$, για την εύρεση του πίνακα $L = (l_{ij})$ με $l_{ij} = 0$ για $i < j$, έχουμε

$$LL^T = A \Leftrightarrow \begin{cases} l_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2} & , i = 1, 2, \dots, n \\ l_{ij} = \frac{1}{l_{jj}} (a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}) & , i = j + 1, \dots, n \end{cases} \quad (1)$$

Η παράσταση $a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2$ είναι πάντοτε θετική, όταν ο πίνακας A είναι θετικά ορισμένος.

Ομοίως, ένας ερμιτιανός ($A^* = \bar{A}^T = A$) και θετικά ορισμένος πίνακας A μπορεί να γραφεί ως γινόμενο της μορφής $A = LL^*$, όπου L είναι πίνακας τριγωνικός κάτω με στοιχεία στην κύρια διαγώνιο θετικά και L^* είναι ο ερμιτιανός συζυγής του πίνακα L .

Παράδειγμα

Έχουμε τον πίνακα $A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 4 \end{bmatrix}$.

Παρατηρούμε ότι ο πίνακας A είναι συμμετρικός ($A^T = A$).

Θα βρούμε τώρα τις ιδιοτιμές του πίνακα A .

$$|A - \lambda I| = 0 \Leftrightarrow \begin{vmatrix} 4 - \lambda & 0 & 1 \\ 0 & 4 - \lambda & 0 \\ 1 & 0 & 4 - \lambda \end{vmatrix} = 0 \Leftrightarrow \lambda_1 = 5, \lambda_2 = 4, \lambda_3 = 3$$

Επομένως ο πίνακας A έχει θετικές ιδιοτιμές τις 5, 4 και 3 οπότε είναι θετικά ορισμένος.

Εφαρμόζουμε τις σχέσεις (1) για να βρούμε τον πίνακα L και ο πίνακας που προκύπτει είναι ο:

$$L = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 1/2 & 0 & \sqrt{15}/2 \end{bmatrix}$$

Επομένως ο πίνακας A μπορεί να γραφεί στη μορφή $A = LL^T$

$$A = LL^T = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 1/2 & 0 & \sqrt{15}/2 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1/2 \\ 0 & 2 & 0 \\ 0 & 0 & \sqrt{15}/2 \end{bmatrix}.$$

Παράδειγμα

Έχουμε τον πίνακα $A = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 4 & 8 & 10 \\ 2 & 8 & 29 & 22 \\ 1 & 10 & 22 & 42 \end{bmatrix}$.

Παρατηρούμε ότι ο πίνακας A είναι συμμετρικός ($A^T = A$).

Θα υπολογίσουμε τώρα τις κύριες υποορίζουσες του πίνακα A .

$$A_1 = \alpha_{11} = 1 > 0$$

$$A_2 = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} = 4 > 0$$

$$A_3 = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} = 36 > 0$$

$$A_4 = \begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{vmatrix} = 576 > 0$$

Παρατηρούμε ότι $A_j > 0$, για κάθε $j = 1, \dots, 4$ και επομένως ο πίνακας A είναι θετικά ορισμένος.

Εφαρμόζουμε τις σχέσεις (1) για να βρούμε τον πίνακα L και ο πίνακας που προκύπτει είναι ο:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 4 & 3 & 0 \\ 1 & 5 & 0 & 4 \end{bmatrix}$$

Επομένως ο πίνακας A μπορεί να γραφεί στη μορφή $A = LL^T$

$$A = LL^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 2 & 4 & 3 & 0 \\ 1 & 5 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & 4 & 5 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

2.2 Εφαρμογές και σχόλια

Όπως έχει ήδη αναφερθεί η παραγοντοποίηση Cholesky είναι ειδική περίπτωση της LU - παραγοντοποίησης. Χρησιμοποιείται συνήθως για την επίλυση των κανονικών εξισώσεων γραμμικών προβλημάτων ελαχίστων τετραγώνων, στη μέθοδο Monte Carlo και στα φίλτρα του Kalman.

Όσον αφορά την επίλυση των κανονικών εξισώσεων γραμμικών προβλημάτων ελαχίστων τετραγώνων έχουμε τα εξής:

Αν ο $m \times n$ πίνακας A έχει τάξη n , τότε ο συμμετρικός $n \times n$ πίνακας $A^T A$ είναι θετικά ορισμένος οπότε η παραγοντοποίηση Cholesky:

$$A^T A = LL^T$$

μπορεί να χρησιμοποιηθεί για τη λύση του συστήματος:

$$A^T A X = A^T B \quad (\text{σύστημα κανονικών εξισώσεων})$$

που έχει την ίδια λύση με αυτή του προβλήματος των γραμμικών ελαχίστων τετραγώνων $AX \approx B$.

Παράδειγμα (Μέθοδος Κανονικών Εξισώσεων)

Γνωρίζουμε ότι για την προσέγγιση με τη μέθοδο ελαχίστων τετραγώνων 5 σημείων με πολυώνυμο 2^{ου} βαθμού ο πίνακας A είναι τύπου Vandermonde και έχουμε:

$$AX = \begin{bmatrix} 1 & c_1 & c_1^2 \\ 1 & c_2 & c_2^2 \\ 1 & c_3 & c_3^2 \\ 1 & c_4 & c_4^2 \\ 1 & c_5 & c_5^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \approx \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = b$$

Με σημεία τα $(-1, 1)$, $(-0.5, 0.5)$, $(0, 0)$, $(0.5, 0.5)$, $(1, 2)$ λαμβάνουμε:

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -0.5 & 0 & 0.5 & 1 \\ 1 & 0.25 & 0 & 0.25 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & -0.5 & 0.25 \\ 1 & 0 & 0 \\ 1 & 0.5 & 0.25 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 2.5 \\ 0 & 2.5 & 0 \\ 2.5 & 0 & 2.125 \end{bmatrix}$$

$$A^T B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -0.5 & 0 & 0.5 & 1 \\ 1 & 0.25 & 0 & 0.25 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0.5 \\ 0 \\ 0.5 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3.25 \end{bmatrix}$$

Η παραγοντοποίηση Cholesky για τον $A^T A$ δίνει:

$$A^T A = \begin{bmatrix} 5 & 0 & 2.5 \\ 0 & 2.5 & 0 \\ 2.5 & 0 & 2.125 \end{bmatrix} = \begin{bmatrix} 2.236 & 0 & 0 \\ 0 & 1.581 & 0 \\ 1.118 & 0 & 0.935 \end{bmatrix} \begin{bmatrix} 2.236 & 0 & 1.118 \\ 0 & 1.581 & 0 \\ 0 & 0 & 0.935 \end{bmatrix} = LL^T$$

Λύνοντας το κάτω τριγωνικό σύστημα $LY = A^T B$ με προς τα εμπρός αντικατάσταση έχουμε:

$$Y = \begin{bmatrix} 1.789 \\ 0.632 \\ 1.336 \end{bmatrix}$$

Λύνοντας το άνω τριγωνικό σύστημα $L^T X = Y$ με προς τα πίσω αντικατάσταση έχουμε τη λύση των ελαχίστων τετραγώνων:

$$X = \begin{bmatrix} 0.086 \\ 0.400 \\ 1.429 \end{bmatrix}.$$

Παραπάνω αναφερθήκαμε στη μέθοδο Monte Carlo, η οποία είναι μια αριθμητική μέθοδος για την επίλυση μαθηματικών προβλημάτων μέσω προσομείωσης τυχαίων αριθμών. Χρησιμοποιείται δε στην προσομείωση και στην ολοκλήρωση.

Η μέθοδος Cholesky χρησιμοποιείται στη μέθοδο Monte Carlo για την προσομείωση συστημάτων με πολλαπλές συσχετισμένες μεταβλητές. Ο πίνακας συσχέτισης αποσυντίθεται για να δώσει τον κάτω τριγωνικό πίνακα L.

Τέλος, γνωρίζουμε ότι τα φίλτρα του Kalman παρέχουν εκτιμήσεις των αληθινών τιμών των μετρήσεων. Θεωρητικά λοιπόν αποτελούν έναν εκτιμητή, πρακτικά όμως είναι ένα μοναδικό εργαλείο για τον έλεγχο πολύπλοκων δυναμικών διεργασιών (π.χ. η ροή ενός πλημμυρισμένου ποταμού, οι τιμές χρηματιστηριακών μετοχών) ή συστημάτων σε οχήματα, πλοία, αεροσκάφη, δορυφόρους,...

Μέσω των φίλτρων του Kalman είναι δυνατή η πρόβλεψη της δυναμικής κατάστασης ενός συστήματος λαμβάνοντας υπόψη την επίδραση όλων των προηγούμενων παρατηρήσεων.

Όταν το μοντέλο παρατήρησης είναι μη - γραμμικό το unscented Kalman filter (UKF) χρησιμοποιεί συνήθως την παραγοντοποίηση Cholesky για να επιλέξει ένα σύνολο των αποκαλούμενων σημείων sigma. Τα φίλτρα Kalman ανιχνεύουν το μέσο όρο του συστήματος σαν ένα διάνυσμα x μήκους n και τη συνδιασπορά σαν έναν πίνακα P $n \times n$. Ο πίνακας P είναι πάντα θετικά ημιορισμένος και μπορεί να παραγοντοποιηθεί σε γινόμενο LL^T . Οι στήλες του L μπορούν να προστεθούν και να αφαιρεθούν από το μέσο x για να σχηματίσουν ένα σύνολο $2n$ διανυσμάτων που ονομάζονται σημεία sigma. Αυτά τα σημεία sigma καταγράφουν πλήρως το μέσο όρο και τη συνδιασπορά του συστήματος.

3. Η QR - παραγοντοποίηση πίνακα

3.1 Η μέθοδος ορθοκανονικοποίησης Gram-Schmidt

Ορισμός 1: (Γραμμική εξάρτηση διανυσμάτων)

i) Τα στοιχεία x_1, x_2, \dots, x_n , $n \in \mathbb{N}^*$, του διανυσματικού χώρου V πάνω στο σώμα K είναι γραμμικώς εξαρτημένα, αν υπάρχουν συντελεστές $\lambda_1, \lambda_2, \dots, \lambda_n \in K$, όχι όλοι μηδέν, έτσι ώστε να ισχύει:

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = 0 .$$

ii) Τα στοιχεία x_1, x_2, \dots, x_n , $n \in \mathbb{N}^*$, του διανυσματικού χώρου V πάνω στο σώμα K είναι γραμμικώς ανεξάρτητα, αν δεν είναι γραμμικώς εξαρτημένα ή ισοδύναμα, αν ισχύει η συνεπαγωγή

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = 0, \lambda_1, \lambda_2, \dots, \lambda_n \in K \implies \lambda_1 = \lambda_2 = \dots = \lambda_n = 0$$

Ορισμός 2:

Το υποσύνολο $\{u_1, u_2, \dots, u_n\}$ του διανυσματικού χώρου V , πάνω στο σώμα K , είναι μία *βάση* του διανυσματικού χώρου V , αν ισχύουν:

i) Τα u_1, u_2, \dots, u_n είναι γραμμικώς ανεξάρτητα,

ii) $[u_1, u_2, \dots, u_n] = V$, δηλαδή τα u_1, u_2, \dots, u_n παράγουν το χώρο V .

Ορισμός 3:

a) Ένα υποσύνολο $S = \{u_1, u_2, \dots, u_n\}$ του Ευκλείδειου ή ορθομοναδιαίου χώρου V , είναι *ορθοκανονικό σύνολο*, όταν ισχύουν:

$$\|u_i\| = 1, \text{ για κάθε } i = 1, 2, \dots, n \quad \text{και} \\ \langle u_i | u_j \rangle = 0, \text{ για κάθε } u_i, u_j \in S \text{ με } i \neq j,$$

ή ισοδύναμα, όταν ισχύει:

$$\langle u_i | u_j \rangle = \delta_{ij} = \begin{cases} 1, & \text{αν } i = j \\ 0, & \text{αν } i \neq j \end{cases} .$$

b) Ένα υποσύνολο $S = \{u_1, u_2, \dots, u_n\}$ του Ευκλείδειου ή ορθομοναδιαίου χώρου V , είναι *ορθοκανονική βάση* του V , όταν είναι ταυτόχρονα βάση και ορθοκανονικό σύνολο του V .

Αφού παραθέσαμε βασικούς ορισμούς παραπάνω, στο σημείο αυτό εξετάζουμε, όταν έχουμε μία βάση ενός Ευκλείδειου ή ορθομοναδιαίου χώρου, πως είναι δυνατόν να κατασκευάσουμε μέσω αυτής μία νέα βάση, η οποία είναι ορθοκανονική. Η κατασκευή αυτή είναι γνωστή ως «**μέθοδος ορθοκανονικοποίησης των Gram-Schmidt**» και γενικεύεται και σε διανυσματικούς χώρους εσωτερικού γινομένου άπειρης διάστασης. Για την παρουσίαση αυτής της μεθόδου είναι αναγκαίο το ακόλουθο λήμμα.

Λήμμα:

Έστω V Ευκλείδειος ή ορθομοναδιαίος χώρος και το ορθοκανονικό υποσύνολό του $S = \{u_1, u_2, \dots, u_n\}$. Τότε ισχύουν

α) Το σύνολο $S = \{u_1, u_2, \dots, u_n\}$ είναι γραμμικώς ανεξάρτητο.

β) Για κάθε $v \in V$ το διάνυσμα

$$u := v - \langle v|u_1 \rangle u_1 - \dots - \langle v|u_n \rangle u_n = v - \sum_{i=1}^n \langle v|u_i \rangle u_i$$

είναι ορθογώνιο προς καθένα από τα u_1, u_2, \dots, u_n .

Θεώρημα:

Κάθε Ευκλείδειος ή ορθομοναδιαίος χώρος V έχει μία τουλάχιστον ορθοκανονική βάση. Αναλυτικότερα, αν $\{u_1, u_2, \dots, u_n\}$ είναι μία βάση του V , τότε το σύνολο $\{\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n\}$, που ορίζεται επαγωγικά ως εξής:

$$\tilde{u}_1 := \frac{u'_1}{\|u'_1\|}, \text{ όπου } u'_1 := u_1,$$

$$\tilde{u}_2 := \frac{u'_2}{\|u'_2\|}, \text{ όπου } u'_2 := u_2 - \langle u_2|\tilde{u}_1 \rangle \tilde{u}_1$$

$$\tilde{u}_3 := \frac{u'_3}{\|u'_3\|}, \text{ όπου } u'_3 := u_3 - \langle u_3|\tilde{u}_1 \rangle \tilde{u}_1 - \langle u_3|\tilde{u}_2 \rangle \tilde{u}_2$$

... ..

$$\tilde{u}_n := \frac{u'_n}{\|u'_n\|}, \text{ όπου } u'_n := u_n - \langle u_n|\tilde{u}_1 \rangle \tilde{u}_1 - \langle u_n|\tilde{u}_2 \rangle \tilde{u}_2 - \dots - \langle u_n|\tilde{u}_{n-1} \rangle \tilde{u}_{n-1},$$

είναι μία ορθοκανονική βάση του Ευκλείδειου ή ορθομοναδιαίου χώρου V .

Παράδειγμα (Μέθοδος ορθοκανονικοποίησης Gram-Schmidt στον \mathbb{R}^3)

Χρησιμοποιώντας τη μέθοδο Gram-Schmidt, από τη βάση

$$\{u_1 = (1,2,0), u_2 = (1,0,1), u_3 = (2,3,1)\},$$

να κατασκευάσετε μία ορθοκανονική βάση του \mathbb{R}^3 .

Λύση Θεωρούμε τον Ευκλείδειο χώρο \mathbb{R}^3 με το κανονικό εσωτερικό γινόμενο και την αντίστοιχη νόρμα που επάγεται από αυτό.

Έχουμε $\|u_1\| = \sqrt{5}$, οπότε είναι $\tilde{u}_1 := \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{5}}(1, 2, 0)$.

Στη συνέχεια έχουμε

$$u'_2 := u_2 - \langle u_2 | \tilde{u}_1 \rangle \tilde{u}_1 = (1,0,1) - \frac{1}{5}(1, 2, 0) = \left(\frac{4}{5}, -\frac{2}{5}, 1\right), \text{ με } \|u'_2\| = \sqrt{\frac{45}{25}} = \frac{3}{\sqrt{5}},$$

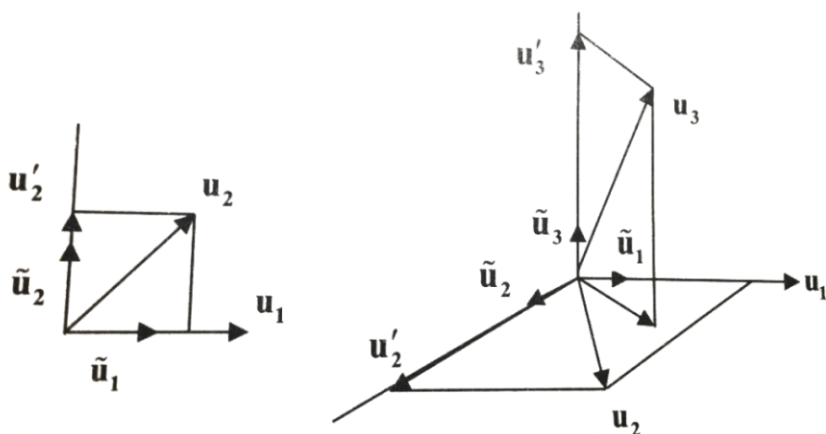
οπότε $\tilde{u}_2 := \frac{u'_2}{\|u'_2\|} = \frac{\sqrt{5}}{3} \left(\frac{4}{5}, -\frac{2}{5}, 1\right) = \frac{1}{3\sqrt{5}}(4, -2, 5)$.

Επίσης είναι

$$u'_3 := u_3 - \langle u_3 | \tilde{u}_1 \rangle \tilde{u}_1 - \langle u_3 | \tilde{u}_2 \rangle \tilde{u}_2 = \frac{1}{45}(-10, 5, 10) = \left(-\frac{2}{9}, \frac{1}{9}, \frac{2}{9}\right)$$

με $\|u'_3\| = \frac{1}{3}$, οπότε τελικά είναι: $\tilde{u}_3 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{2}{3}\right)$.

Στα σχήματα που ακολουθούν φαίνεται γεωμετρικά η διαδικασία ορθοκανονικοποίησης διανυσμάτων για $n = 2$ και $n = 3$.



3.2 Ο αλγόριθμος της QR παραγοντοποίησης

Θεωρούμε τον πίνακα $A \in M_{m \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} του οποίου τα διανύσματα στήλες είναι γραμμικώς ανεξάρτητα (δηλαδή ο βαθμός του πίνακα A είναι $\rho = n$, γενικά $\rho \leq n$). Υποθέτουμε ότι, τα ορθοκανονικοποιημένα διανύσματα q_1, q_2, \dots, q_n που λαμβάνονται με τη μέθοδο ορθοκανονικοποίησης Gram-Schmidt, είναι οι στήλες ενός πίνακα Q , δηλαδή

$$Q = [q_1 \ q_2 \ \dots \ q_n].$$

Θα προσπαθήσουμε να βρούμε τη σχέση που ισχύει μεταξύ των πινάκων A και Q . Αν u_1, u_2, \dots, u_n είναι τα γραμμικώς ανεξάρτητα διανύσματα – στήλες του πίνακα A , τότε γράφουμε $A = [u_1 \ u_2 \ \dots \ u_n]$.

Επειδή η βάση $\{q_1, q_2, \dots, q_n\}$ του διανυσματικού χώρου $C_S(A)$ είναι ορθοκανονική, γνωρίζουμε ότι ισχύουν:

$$\begin{cases} u_1 = \langle u_1 | q_1 \rangle q_1 + \langle u_1 | q_2 \rangle q_2 + \dots + \langle u_1 | q_n \rangle q_n \\ u_2 = \langle u_2 | q_1 \rangle q_1 + \langle u_2 | q_2 \rangle q_2 + \dots + \langle u_2 | q_n \rangle q_n \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ u_n = \langle u_n | q_1 \rangle q_1 + \langle u_n | q_2 \rangle q_2 + \dots + \langle u_n | q_n \rangle q_n \end{cases}$$

ή ισοδύναμα

$$[u_1 \ u_2 \ \dots \ u_n] = [q_1 \ q_2 \ \dots \ q_n] \begin{bmatrix} \langle u_1 | q_1 \rangle & \langle u_2 | q_1 \rangle & \dots & \langle u_n | q_1 \rangle \\ \langle u_1 | q_2 \rangle & \langle u_2 | q_2 \rangle & \dots & \langle u_n | q_2 \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle u_1 | q_n \rangle & \langle u_2 | q_n \rangle & \dots & \langle u_n | q_n \rangle \end{bmatrix}$$

$$\Leftrightarrow A = QR,$$

όπου ο πίνακας R είναι τριγωνικός πάνω, αφού από τη μέθοδο Gram-Schmidt είναι γνωστό ότι $\langle u_i | q_j \rangle = 0$, για κάθε $i=1,2,\dots,j-1$.

Άρα είναι

$$R = \begin{bmatrix} \langle u_1 | q_1 \rangle & \langle u_2 | q_1 \rangle & \dots & \langle u_n | q_1 \rangle \\ 0 & \langle u_2 | q_2 \rangle & \dots & \langle u_n | q_2 \rangle \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & \langle u_n | q_n \rangle \end{bmatrix}.$$

Έτσι έχουμε αποδείξει το θεώρημα που ακολουθεί:

Θεώρημα:

Αν A είναι ένας $m \times n$ πίνακας με στήλες γραμμικώς ανεξάρτητα διανύσματα, τότε ο A μπορεί να γραφεί ως γινόμενο πινάκων της μορφής

$$A = QR,$$

όπου Q ο $m \times n$ πίνακας με στήλες ορθοκανονικά διανύσματα και R είναι ένας $n \times n$ αντιστρέψιμος τριγωνικός πάνω πίνακας.

Παράδειγμα

Να βρείτε την QR – παραγοντοποίηση του πίνακα

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 1 & 2 & 0 \end{bmatrix}$$

Παρατηρούμε ότι οι στήλες του πίνακα A είναι γραμμικώς ανεξάρτητα διανύσματα. Στη συνέχεια με τη μέθοδο Gram-Schmidt έχουμε:

$$u_1 = (1, 0, 1) = u_1', \quad \|u_1'\| = \sqrt{2} \quad \text{και} \quad q_1 = \frac{u_1'}{\|u_1'\|} = \frac{1}{\sqrt{2}}(1, 0, 1)$$

$$u_2' = u_2 - \langle u_2 | q_1 \rangle q_1 = (0, 1, 2) - \frac{2}{\sqrt{2}} \frac{1}{\sqrt{2}}(1, 0, 1) = (0, 1, 2) - (1, 0, 1) = (-1, 1, 1)$$

οπότε θα είναι

$$\|u_2'\| = \sqrt{3} \quad \text{και} \quad q_2 = \frac{u_2'}{\|u_2'\|} = \frac{1}{\sqrt{3}}(-1, 1, 1).$$

Επίσης έχουμε

$$\begin{aligned} u_3' &= u_3 - \langle u_3 | q_1 \rangle q_1 - \langle u_3 | q_2 \rangle q_2 = (2, 1, 0) - \frac{2}{\sqrt{2}} \frac{1}{\sqrt{2}}(1, 0, 1) - \left(-\frac{1}{\sqrt{3}}\right) \frac{1}{\sqrt{3}}(-1, 1, 1) \\ &= (2, 1, 0) - (1, 0, 1) + \left(-\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = \left(\frac{2}{3}, \frac{4}{3}, -\frac{2}{3}\right) \end{aligned}$$

οπότε θα είναι

$$\|u'_3\| = \sqrt{\frac{24}{9}} = \frac{2\sqrt{6}}{3} \quad \text{και}$$

$$q_3 = \frac{u'_3}{\|u'_3\|} = \frac{1}{\frac{2\sqrt{6}}{3}} \left(\frac{2}{3}, \frac{4}{3}, -\frac{2}{3} \right) = \frac{\sqrt{6}}{12} (2, 4, -2) = \frac{\sqrt{6}}{6} (1, 2, -1)$$

Άρα έχουμε

$$Q = [q_1 \ q_2 \ q_3] = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{\sqrt{6}}{6} \\ 0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{6}}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{\sqrt{6}}{6} \end{bmatrix}$$

$$R = \begin{bmatrix} \langle u_1 | q_1 \rangle & \langle u_2 | q_1 \rangle & \langle u_3 | q_1 \rangle \\ 0 & \langle u_2 | q_2 \rangle & \langle u_3 | q_2 \rangle \\ 0 & 0 & \langle u_3 | q_3 \rangle \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} & -\frac{1}{\sqrt{3}} \\ 0 & 0 & \frac{4\sqrt{6}}{6} \end{bmatrix}$$

Επομένως, $A = QR$.

3.3 Εφαρμογές και σχόλια

Αν ένας αντιστρέψιμος πίνακας $A \in M_{n \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} έχει παραγοντοποιηθεί στη μορφή QR , $A = QR$, τότε κάθε σύστημα $AX = B \Leftrightarrow QRX = B$ επιλύεται εύκολα σε δύο βήματα. Αρχικά, θεωρούμε το σύστημα $QY = B$ που έχει λύση $Y = Q^*B$ και στη συνέχεια λύνουμε το τριγωνικό σύστημα $RX = Y$. Επιπλέον, ο αντίστροφος του A , $A^{-1} = (QR)^{-1} = R^{-1}Q^*$, είναι άμεσα υπολογίσιμος αφού ο R^{-1} προκύπτει από την επίλυση της εξίσωσης $RR^{-1} = I_n$, η οποία είναι ευκόλως επιλύσιμη καθώς ο R είναι ένας $n \times n$ αντιστρέψιμος τριγωνικός πάνω πίνακας.

Επιπλέον, το πρόβλημα των ελαχίστων τετραγώνων διατυπώνεται ως εξής:

Δεδομένου ενός πίνακα $A \in M_{m \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} και ενός διανύσματος $B \in \mathbb{R}^m$ ή \mathbb{C}^m να υπολογιστεί ένα $x \in \mathbb{R}^n$ ή \mathbb{C}^n τέτοιο ώστε η απόσταση $\|AX - B\|_2$ να είναι η ελάχιστη δυνατή. Όταν ο βαθμός του πίνακα A είναι $p = n$, μια δημοφιλής μέθοδος επίλυσης του προβλήματος αυτού είναι μέσω της παραγοντοποίησης QR του πίνακα A , $A = QR$. Συγκεκριμένα έχουμε

$$AX = B \Leftrightarrow QRX = B \Leftrightarrow X = R^{-1}Q^*B$$

αφού ο R είναι ένας $n \times n$ αντιστρέψιμος τριγωνικός πάνω πίνακας. Επομένως η λύση του προβλήματος ελαχίστων τετραγώνων $\min_{x \in \mathbb{C}^n} \|AX - B\|_2$ είναι $x = R^{-1}Q^*B$.

Τέλος, η παραγοντοποίηση QR είναι χρήσιμη για τον υπολογισμό δυνάμεων του πίνακα A . Ακόμη, αν $A \in M_{n \times n}(K)$, $K = \mathbb{R}$ ή \mathbb{C} η παραγοντοποίηση QR βοηθάει στην προσέγγιση των ιδιοτιμών και των ιδιοδιανυσμάτων του πίνακα A καθώς ισχύουν τα εξής:

$$\mathbb{R}: Q^T Q = I \Rightarrow Q^T = Q^{-1} \text{ (ο πίνακας } Q \text{ είναι ορθογώνιος)}$$

$$\mathbb{C}: Q^* Q = I \Rightarrow Q^* = Q^{-1} \text{ (ο πίνακας } Q \text{ είναι ορθομοναδιαίος)}$$

με αποτέλεσμα η εφαρμογή μετασχηματισμού ομοιότητας με τον Q να μην απαιτεί αντιστροφή. (γνωρίζουμε πως όμοιοι πίνακες έχουν τις ίδιες ακριβώς ιδιοτιμές). Στην περίπτωση μας ο πίνακας R που βρίσκουμε με την παραγοντοποίηση QR δεν έχει τις ίδιες ιδιοτιμές με τον A , όμως ο πίνακας $A^{(1)} = Q^T A Q$ (όταν $K = \mathbb{R}$) είναι όμοιος με τον A και άρα έχει τις ίδιες ιδιοτιμές. Εφαρμόζοντας τον παραπάνω μετασχηματισμό ομοιότητας επαναληπτικά παίρνουμε τον αλγόριθμο $A^{(k)} = Q^{(k)} R^{(k)}$, $A^{(k+1)} = (Q^{(k)})^T A^{(k)} Q^{(k)}$. Αποδεικνύεται ότι $A^{(k \rightarrow \infty)}$ είναι ένας «block» άνω τριγωνικός πίνακας όπου οι πίνακες της κύριας διαγωνίου είναι 1×1 ή 2×2 . Για την πρώτη περίπτωση, τα διαγώνια στοιχεία είναι οι ιδιοτιμές του A . Για τη δεύτερη περίπτωση, οι 2×2 υποπίνακες έχουν δύο ιδιοτιμές που είναι συζυγείς μιγαδικές (για A πραγματικούς) και είναι ιδιοτιμές του A .

4. Η SVD - παραγοντοποίηση πίνακα

4.1 Ανάλυση ιδιοζουσών τιμών

Θεωρούμε πίνακα $A \in M_{m \times n}(\mathbb{R})$. Τότε ο πίνακας $A^T A$ είναι συμμετρικός θετικά ημιορισμένος $n \times n$ πίνακας, οπότε υπάρχει ορθοκανονική βάση $\{u_1, u_2, \dots, u_n\}$ ιδιοδιανυσμάτων αυτού. Αν υποθέσουμε ότι

$$A^T A u_j = \lambda_j u_j, \quad j = 1, 2, \dots, n,$$

όπου οι ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_n$ δεν είναι κατ' ανάγκη διακεκριμένες, τότε ως προς το κανονικό εσωτερικό γινόμενο του \mathbb{R}^n έχουμε

$$\langle A u_i | A u_j \rangle = (A u_i)^T I_n A u_j = u_i^T A^T A u_j = \lambda_j u_i^T u_j = \lambda_j \delta_{ij} = \begin{cases} \lambda_j, & i = j \\ 0, & i \neq j \end{cases}.$$

Επομένως τα διανύσματα $A u_1, A u_2, \dots, A u_n$ είναι ανά δύο ορθογώνια και έχουν μέτρο $\|A u_i\|^2 = \lambda_i, i = 1, 2, \dots, n$. Επειδή ο πίνακας $A^T A$ είναι θετικά ημιορισμένος οι ιδιοτιμές του είναι μη αρνητικές, οπότε ισχύει

$$\|A u_i\| = \sqrt{\lambda_i}, \quad i = 1, 2, \dots, n.$$

Ορισμός:

Ιδιάζουσες τιμές ενός πίνακα $A \in M_{m \times n}(\mathbb{R})$ είναι οι αριθμοί

$$\sigma_i = \sqrt{\lambda_i}, \quad i = 1, 2, \dots, n,$$

όπου $\lambda_i, i = 1, 2, \dots, n$ είναι οι ιδιοτιμές του πίνακα $A^T A$. Διατάσσουμε συνήθως τις ιδιάζουσες τιμές σε φθίνουσα σειρά $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.

Θεώρημα:

Έστω $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ γραμμική απεικόνιση με τύπο $f(r) = Ar$, όπου $A \in M_{m \times n}(\mathbb{R})$. Τότε ισχύουν:

- Υπάρχει ορθοκανονική βάση $\{u_1, u_2, \dots, u_n\}$ του \mathbb{R}^n , έτσι ώστε τα διανύσματα $f(u_i) = A u_i, i = 1, 2, \dots, n$ είναι ορθογώνια ανά δύο και έχουν μέτρο $\|A u_i\| = \sigma_i = \sqrt{\lambda_i}, i = 1, 2, \dots, n$.
- Αν ο βαθμός του πίνακα A , $\rho(A) = r$, τότε είναι $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ και $\sigma_{r+1} = \dots = \sigma_n = 0$.

Ας θεωρήσουμε τώρα τα μοναδιαία και ορθογώνια ανά δύο διανύσματα

$u_1 = \frac{1}{\sigma_1} Au_1, u_2 = \frac{1}{\sigma_2} Au_2, \dots, u_r = \frac{1}{\sigma_r} Au_r$, όπου u_1, u_2, \dots, u_r είναι τα ιδιοδιανύσματα του $A^T A$ που αντιστοιχούν στις μη μηδενικές ιδιοτιμές του με φθίνουσα σειρά. Τότε αυτά επεκτείνονται σε μία ορθοκανονική βάση του \mathbb{R}^m , έστω $\{u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_m\}$, για την οποία ισχύουν

$$Au_i = \sigma_i u_i, i = 1, 2, \dots, r,$$

$$Au_i = 0, i = r + 1, \dots, n.$$

Αν $V = [u_1 \ u_2 \ \dots \ u_n]$ είναι ο πίνακας με στήλες τα στοιχεία της βάσης $\{u_1, u_2, \dots, u_n\}$, τότε θα έχουμε

$$AV = A[u_1 \ u_2 \ \dots \ u_n] = [Au_1 \ Au_2 \ \dots \ Au_n] = [\sigma_1 u_1 \ \dots \ \sigma_r u_r \ 0 \ \dots \ 0]$$

$$= [u_1, u_2, \dots, u_r, u_{r+1}, \dots, u_m] \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & \dots & \sigma_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & & \vdots & \vdots & \ddots \end{bmatrix} = U\Sigma,$$

όπου $U = [u_1, \dots, u_r, u_{r+1}, \dots, u_m]$. Επειδή από τον ορισμό τους οι πίνακες U και V είναι ορθογώνιοι (δηλαδή $U^T = U^{-1}$ και $V^T = V^{-1}$), έχουμε

$$A = U\Sigma V^{-1} = U\Sigma V^T.$$

Με απευθείας υπολογισμό του γινομένου $U\Sigma V^T$ έχουμε την ισότητα

$$A = \sigma_1 u_1 u_1^T + \sigma_2 u_2 u_2^T + \dots + \sigma_r u_r u_r^T.$$

Έτσι έχουμε αποδείξει το ακόλουθο πολύ σημαντικό θεώρημα:

Θεώρημα:

Κάθε πίνακας $A \in M_{m \times n}(\mathbb{R})$ με $\rho(A) = r$ γράφεται ως γινόμενο

$$A = U\Sigma V^T, \quad (1)$$

όπου U είναι ένας $m \times m$ ορθογώνιος πίνακας, V είναι ένας $n \times n$ ορθογώνιος πίνακας και $\Sigma = (\sigma_{ij})$ είναι ένας $m \times n$ πίνακας με $\sigma_{ii} = \sigma_i$, $i = 1, 2, \dots, r$, $\sigma_{ii} = 0$, για $i = r+1, \dots, m$ και $\sigma_{ij} = 0$, για $i \neq j$.

Επιπλέον ο πίνακας A γράφεται στη μορφή

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_r u_r v_r^T \quad (2)$$

όπου u_i και v_i είναι οι αντίστοιχες στήλες των πινάκων U και V .

Παράδειγμα

Έστω πίνακας $A = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$.

α) Να βρείτε τις ιδιάζουσες τιμές του A .

β) Να κατασκευάσετε μια παραγοντοποίηση SVD του πίνακα A .

Λύση α) Θα βρω τις ιδιοτιμές του πίνακα

$$A^T A = \begin{bmatrix} 2 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$$

$$|A^T A - \lambda I| = 0 \Leftrightarrow \begin{vmatrix} 5 - \lambda & 3 \\ 3 & 5 - \lambda \end{vmatrix} = 0 \Leftrightarrow \lambda_1 = 8, \lambda_2 = 2.$$

Οπότε οι ιδιάζουσες τιμές του A είναι

$$\sigma_1 = \sqrt{\lambda_1} = 2\sqrt{2} \quad \text{και} \quad \sigma_2 = \sqrt{\lambda_2} = \sqrt{2}.$$

β) Θα βρω τα ιδιοδιανύσματα που αντιστοιχούν στις ιδιοτιμές $\lambda_1 = 8$ και $\lambda_2 = 2$.

$\lambda_1 = 8$:

$$\begin{bmatrix} 8 - 5 & -3 \\ -3 & 8 - 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow y = x$$

Άρα έχω το ιδιοδιάνυσμα $x_1 = (1, 1)$.

$\lambda_2 = 2$:

$$\begin{bmatrix} 2-5 & -3 \\ -3 & 2-5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow y = -x$$

Άρα έχω το ιδιοδιάνυσμα $x_2 = (-1,1)$ που είναι ορθογώνιο με το x_1 .

Οπότε θα είναι

$$u_1 = \frac{1}{\sqrt{2}} (1,1) \text{ και } u_2 = \frac{1}{\sqrt{2}} (-1,1).$$

Τότε θα είναι

$$Au_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ 0 \end{bmatrix}, \text{ με } \|Au_1\| = 2\sqrt{2} = \sigma_1,$$

$$Au_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{2} \end{bmatrix}, \text{ με } \|Au_2\| = \sqrt{2} = \sigma_2.$$

$$\text{Ισχύει } u_1 = \frac{1}{\sigma_1} Au_1 = (1,0) \text{ και } u_2 = \frac{1}{\sigma_2} Au_2 = (0,1)$$

Επομένως

$$A = U\Sigma V^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

Συνεχίζοντας, όσον αφορά το σώμα των μιγαδικών ισχύει το ακόλουθο θεώρημα:

Θεώρημα:

Αν $A \in M_{m \times n}(\mathbb{C})$ με $\rho(A) = r$ ένας $m \times n$ μιγαδικός πίνακας τότε υπάρχουν ορθομοναδιαίοι πίνακες $U \in \mathbb{C}^{m \times m}$ και $V \in \mathbb{C}^{n \times n}$ τέτοιοι ώστε

$$A = U \Sigma V^* \quad (1)$$

όπου $\Sigma = (\sigma_{ij})$ είναι ένας $m \times n$ πίνακας με $\sigma_{ii} = \sigma_i$, $i = 1, 2, \dots, r$, $\sigma_{ii} = 0$, για $i = r+1, \dots, m$ και $\sigma_{ij} = 0$, για $i \neq j$.

Επιπλέον $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min\{m,n\}} \geq 0$.

Έχουμε τους θετικά ημιορισμένους ερμιτιανούς πίνακες AA^* και A^*A .

- i. Οι μη μηδενικές ιδιάζουσες τιμές του A είναι οι τετραγωνικές ρίζες των μη μηδενικών ιδιοτιμών των ερμιτιανών πινάκων A^*A και AA^* .
- ii. Τα αριστερά ιδιάζοντα διανύσματα του A , δηλαδή τα διανύσματα u_i του πίνακα U , είναι τα ιδιοδιανύσματα του πίνακα AA^* .
- iii. Τα δεξιά ιδιάζοντα διανύσματα του A , δηλαδή τα διανύσματα v_i του πίνακα V , είναι τα ιδιοδιανύσματα του πίνακα A^*A .

Επιπλέον έχουμε ότι ο πίνακας A γράφεται ως άθροισμα r πινάκων βαθμού 1.

$$A = \sum_{i=1}^r \sigma_i u_i v_i^* \quad (2)$$

Η συγκεκριμένη γραφή οδηγεί σε προσέγγιση του πίνακα A από πίνακες μικρότερου του $r = \text{βαθμός}(A)$ βαθμού.

4.2 Ευστάθεια Τετραγωνικών Γραμμικών Συστημάτων

Έστω ένας $n \times n$ αντιστρέψιμος πίνακας A με παραγοντοποίηση SVD

$$A = UV^* = \sum_{i=1}^n \sigma_i u_i v_i^*$$

Τότε ισχύει:

$$(U \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\} V^*) (V \text{diag}\{\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}\} U^*) = I_n$$

Δηλαδή

$$A^{-1} = V \text{diag}\{\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}\} U^*$$

Συνεπώς το σύστημα $AX = B$ έχει λύση

$$X = A^{-1}B = \sum_{i=1}^n \left(\frac{u_i^* B}{\sigma_i} \right) v_i$$

Η γραφή αυτή δείχνει ότι μικρές διαταραχές στα στοιχεία του A ή του B μπορούν να προκαλέσουν σχετικά μεγάλες διαταραχές στη λύση X αν η ιδιάζουσα τιμή σ_n είναι “μικρή”, δηλαδή αν ο A είναι “κοντά” σε μη αντιστρέψιμο πίνακα.

Υπενθύμιση: Ο αντίστροφος ενός πίνακα $A \in M_n(K)$ υπάρχει αν και μόνο αν η ορίζουσά του είναι διάφορη του μηδέν, δηλαδή ισχύει:

$$A \text{ αντιστρέψιμος} \Leftrightarrow \det A \neq 0 \quad .$$

4.3 Ψευδοαντίστροφος Πίνακας

Όπως είδαμε στην προηγούμενη παράγραφο, αν ο πίνακας $A \in \mathbb{C}^{n \times n}$ είναι αντιστρέψιμος με παραγοντοποίηση SVD της μορφής $A=UV^*=U\text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\}V^*$ ($\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$) τότε

$$A^{-1} = V\text{diag}\{\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}\}U^*$$

Στην περίπτωση που ο A είναι ένας $m \times n$ πίνακας βαθμού $\rho(A) = r$ και με παραγοντοποίηση SVD

$$A = U \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & & \vdots & \vdots & \ddots \end{bmatrix}_{(m \times n)} V^*$$

ορίζουμε τον $n \times m$ (Moore-Penrose) ψευδοαντίστροφο του A ,

$$A^\dagger = V \begin{bmatrix} \sigma_1^{-1} & 0 & \dots & 0 & 0 & \dots \\ 0 & \sigma_2^{-1} & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_r^{-1} & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & & \vdots & \vdots & \ddots \end{bmatrix}_{(n \times m)} U^*$$

Ο ψευδοαντίστροφος A^\dagger ικανοποιεί τις σχέσεις:

1. $AA^\dagger A = A$
2. $A^\dagger A A^\dagger = A^\dagger$
3. Οι πίνακες AA^\dagger και $A^\dagger A$ είναι ερμιτιανοί.

Αν ο πίνακας A είναι $m \times n$ ($m > n$) με βαθμό $\rho(A) = n$, τότε το πρόβλημα ελαχίστων τετραγώνων $\min_{x \in \mathbb{C}^n} \|AX - B\|_2$ (για το οποίο έγινε λόγος σε προηγούμενο κεφάλαιο) έχει λύση

$$X = R^{-1}Q^*B$$

όπου Q και R είναι οι πίνακες της παραγοντοποίησης QR.

Εύκολα επαληθεύεται ότι:

$$X = R^{-1}Q^*B \Leftrightarrow QRX = QQ^*B = B \text{ (αφού } Q^* = Q^{-1}, Q \text{ ορθομοναδιαίος)}$$

$$\Leftrightarrow (R^*Q^*)(QR)X = R^*Q^*B$$

$$\Leftrightarrow A^*AX = A^*B$$

$$\Leftrightarrow X = (A^*A)^{-1}A^*B$$

$$\Leftrightarrow X = A^\dagger B$$

Παράδειγμα

Κατασκευάστε μια παραγοντοποίηση SVD του πίνακα $A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & -1 & 1 \end{bmatrix}$ και του A^\dagger .

Λύση

Υπενθύμιση: οι πίνακες AA^* και A^*A έχουν κοινές μη αρνητικές ιδιοτιμές $\lambda_1, \lambda_2, \dots$ των οποίων οι τετραγωνικές ρίζες είναι οι ιδιάζουσες τιμές του πίνακα A (οι μη μηδενικές). Τα ιδιοδιανύσματα του AA^* είναι οι στήλες u_1, u_2, \dots του πίνακα U , ενώ τα ιδιοδιανύσματα του A^*A είναι οι στήλες v_1, v_2, \dots του πίνακα V .

Θεωρούμε τους πίνακες

$$AA^* = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$A^*A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & 2 & -1 & 1 \\ 0 & -1 & 2 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$$

Υπολογίζουμε τις ιδιοτιμές του AA^*

$$|\lambda I - AA^*| = 0 \Leftrightarrow \begin{vmatrix} \lambda - 2 & -1 & 0 \\ -1 & \lambda - 2 & 0 \\ 0 & 0 & \lambda - 4 \end{vmatrix} = 0 \Leftrightarrow \lambda_1 = 4, \lambda_2 = 3, \lambda_3 = 1.$$

Οπότε οι μη μηδενικές ιδιάζουσες τιμές του A είναι

$$\sigma_1 = \sqrt{\lambda_1} = 2, \sigma_2 = \sqrt{\lambda_2} = \sqrt{3} \text{ και } \sigma_3 = \sqrt{\lambda_3} = 1.$$

Υπολογίζουμε τα u_1, u_2, \dots ιδιοδιανύσματα του πίνακα A^*A .

$\lambda_1 = 4$:

$$\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 2 & 1 & -1 \\ 0 & 1 & 2 & 1 \\ -1 & -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow (x, y, z, w) = w(1, 1, -1, 1)$$

Άρα έχω το ιδιοδιάνυσμα $x_1 = (1, 1, -1, 1)$.

Οπότε κανονικοποιώντας έχουμε

$$u_1 = \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2} \right)$$

$\lambda_2 = 3$:

$$\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ -1 & -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow (x, y, z, w) = z(2, -1, 1, 0)$$

Άρα έχω το ιδιοδιάνυσμα $x_2 = (2, -1, 1, 0)$ που είναι ορθογώνιο με το x_1 .

Οπότε κανονικοποιώντας έχουμε

$$u_2 = \frac{1}{\sqrt{6}} (2, -1, 1, 0)$$

$\lambda_3 = 1$:

$$\begin{bmatrix} -2 & 0 & 0 & -1 \\ 0 & -1 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ -1 & -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow (x,y,z,w) = z(0,1,1,0)$$

Άρα έχω το ιδιοδιάνυσμα $x_3 = (0,1,1,0)$ που είναι ορθογώνιο με το x_1 και το x_2 .

Οπότε κανονικοποιώντας έχουμε

$$u_3 = \frac{1}{\sqrt{2}}(0,1,1,0)$$

Μας περισσεύει μια ιδιοτιμή η οποία είναι η $\lambda_4 = 0$ για την οποία έχουμε:

$$\begin{bmatrix} -3 & 0 & 0 & -1 \\ 0 & -2 & 1 & -1 \\ 0 & 1 & -2 & 1 \\ -1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow (x,y,z,w) = x(1,1,-1,-3)$$

Άρα έχω το ιδιοδιάνυσμα $x_4 = (1,1,-1,-3)$ που είναι ορθογώνιο με το x_1 , το x_2 και το x_3 .

Οπότε κανονικοποιώντας έχουμε

$$u_4 = \frac{1}{\sqrt{12}}(1,1,-1,-3)$$

Υπολογίζουμε τέλος τα u_1, u_2, \dots ιδιοδιανύσματα του πίνακα $AA^* = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

$\lambda_1 = 4$:

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow (x,y,z) = z(0,0,1)$$

Άρα έχω το ιδιοδιάνυσμα $\gamma_1 = (0,0,1)$.

Οπότε κανονικοποιώντας έχουμε

$$u_1 = (0,0,1)$$

$\lambda_2 = 3$:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow (x,y,z) = x(1,1,0)$$

Άρα έχω το ιδιοδιάνυσμα $\gamma_2 = (1,1,0)$ που είναι ορθογώνιο με το γ_1 .

Οπότε κανονικοποιώντας έχουμε

$$u_2 = \frac{1}{\sqrt{2}}(1,1,0)$$

$\lambda_3 = 1$:

$$\begin{bmatrix} -1 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Leftrightarrow (x,y,z) = x(1,-1,0)$$

Άρα έχω το ιδιοδιάνυσμα $\gamma_3 = (1,-1,0)$ που είναι ορθογώνιο με το γ_1 και το γ_2 .

Οπότε κανονικοποιώντας έχουμε

$$u_3 = \frac{1}{\sqrt{2}}(1,-1,0)$$

Άρα

$$A = U\Sigma V^* = [u_1 \ u_2 \ u_3] \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \\ v_3^* \\ v_4^* \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & -\frac{1}{\sqrt{12}} & -\frac{3}{\sqrt{12}} \end{bmatrix}$$

$$= \sigma_1 u_1 u_1^* + \sigma_2 u_2 u_2^* + \sigma_3 u_3 u_3^*$$

$$= 2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + \sqrt{3} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & 0 \end{bmatrix} + 1 \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

Και για τον ψευδοαντίστροφο A^\dagger

$$A^\dagger = V \Sigma^\dagger U^* = [u_1 \ u_2 \ u_3 \ u_4] \begin{bmatrix} \sigma_1^{-1} & 0 & 0 \\ 0 & \sigma_2^{-1} & 0 \\ 0 & 0 & \sigma_3^{-1} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1^* \\ u_2^* \\ u_3^* \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} & \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{12}} \\ \frac{1}{2} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{12}} \\ -\frac{1}{2} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{12}} \\ \frac{1}{2} & 0 & 0 & -\frac{3}{\sqrt{12}} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

$$= \sigma_1^{-1} u_1 u_1^* + \sigma_2^{-1} u_2 u_2^* + \sigma_3^{-1} u_3 u_3^*$$

$$= \frac{1}{2} \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} + \frac{1}{\sqrt{3}} \begin{bmatrix} \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} + 1 \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} .$$

4.4 Εφαρμογές και σχόλια

Η ανάλυση ιδιζουσών τιμών ενός πίνακα και η παραγοντοποίηση SVD είναι ένα αξιολογικό μαθηματικό “τέχνασμα” με ποικίλες εφαρμογές όπως αναφέρονται παρακάτω:

a) Υπολογισμός του βαθμού ενός πίνακα:

Βαθμός πίνακα A := ο αριθμός που εκφράζει το μέγιστο πλήθος των γραμμικά ανεξάρτητων στηλών (ή ισοδύναμα γραμμών) του πίνακα A . Η σ_r είναι η μικρότερη μη μηδενική ιδιάζουσα τιμή. Έτσι μπορούμε να βρούμε και το βαθμό του πίνακα A καθώς ισχύει:

$$\text{Βαθμός}(A) = \text{Βαθμός}(\Sigma) = r$$

b) Υπολογισμός του αντιστρόφου ενός πίνακα:

Όπως αναλύσαμε σε προηγούμενη παράγραφο, αν A ένας $n \times n$ αντιστρέψιμος πίνακας με παραγοντοποίηση SVD ώστε $A = U\Sigma^*U^*$ τότε

$$A^{-1} = U \text{diag}\{\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}\} U^*$$

c) Υπολογισμός του βαθμού κατάστασης ενός αντιστρέψιμου πίνακα:

Έστω ότι έχουμε το σύστημα $AX = B$. Ο βαθμός κατάστασης “ποσοτικοποιεί” την ευαισθησία του συστήματος και ορίζεται ως το πηλίκο της μεγαλύτερης ως προς τη μικρότερη ιδιάζουσα τιμή του πίνακα A ($\sigma_n(A) > 0$), δηλαδή

$$\kappa(A) = \|A\| \|A^{-1}\| = \frac{\sigma_1(A)}{\sigma_n(A)} .$$

Όσο μεγαλύτερο είναι το πηλίκο τόσο ο πίνακας A είναι πιο κοντά στον να μην είναι αντιστρέψιμος. Οι πίνακες με σχετικά μικρό βαθμό κατάστασης ονομάζονται πίνακες καλής κατάστασης (*well conditioned*), ενώ αν μια μικρή αλλαγή στον πίνακα B μπορεί να προκαλέσει σχετικά μεγάλες αλλαγές στη λύση X , τότε λέμε ότι ο πίνακας A είναι κακής κατάστασης (*ill conditioned*). Το τελευταίο συμβαίνει αν η ιδιάζουσα τιμή σ_n είναι “μικρή”, δηλαδή αν ο A είναι “κοντά” σε μη αντιστρέψιμο πίνακα.

d) Λύση του συστήματος ελαχίστων τετραγώνων:

Όπως αναφερθήκαμε σε προηγούμενη παράγραφο το πρόβλημα ελαχίστων τετραγώνων $\min_{x \in \mathbb{C}^n} \|AX - B\|_2$ έχει λύση $X = R^{-1}Q^*B$, όπου Q και R είναι οι πίνακες της παραγοντοποίησης QR. Αποδείξαμε ότι η λύση μέσω παραγοντοποίησης SVD και με τη βοήθεια του ψευδοαντίστροφου πίνακα A^\dagger είναι η

$$X = A^\dagger B .$$

e) Υπολογισμός του ψευδοαντίστροφου πίνακα A^\dagger :

Γνωρίζουμε ότι αν A είναι ένας $m \times n$ πίνακας βαθμού $\rho(A) = r$, τότε ο ψευδοαντίστροφός του είναι ο

$$A^\dagger = V \begin{bmatrix} \sigma_1^{-1} & 0 & \dots & 0 & 0 & \dots \\ 0 & \sigma_2^{-1} & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & \dots & \sigma_r^{-1} & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & & \vdots & \vdots & \ddots \end{bmatrix}_{(n \times m)} U^* .$$

f) Υπολογισμός της λύσης ομογενών συστημάτων:

Ας υποθέσουμε ότι $B = 0$, τότε το γραμμικό σύστημα λέγεται ομογενές: $AX = 0$.

Υποθέτουμε ότι A ένας $m \times n$ πίνακας και $A = USV^T$.

Η λύση με την ελάχιστη νόρμα είναι $X = 0$ (τετριμμένη λύση).

Για την λύση ομογενών γραμμικών συστημάτων μετασχηματίζω τη λύση βάζοντας τον περιορισμό $\|X\| = 1$. Τότε έχω το πρόβλημα $\min_{\|X\|=1} \|AX\|$.

Ειδική περίπτωση: βαθμός(A) = $n - 1$ ($m \geq n - 1, \sigma_n = 0$) η λύση είναι

$$X = \alpha v_n \text{ (}\alpha \text{ σταθερά)}$$

Γενική περίπτωση: βαθμός(A) = $n - k$ ($m \geq n - k, \sigma_{n-k+1} = \dots = \sigma_n = 0$) η λύση είναι

$$X = \alpha_1 v_{n-k+1} + \alpha_2 v_{n-k} + \dots + \alpha_k v_n \text{ (}\alpha_i \text{ σταθερές)}$$

$$\text{με } \alpha_1^2 + \alpha_2^2 + \dots + \alpha_k^2 = 1 .$$

g) Ελάττωση όγκου και συμπίεση δεδομένων:

Μια σημαντική εφαρμογή της ανάλυσης ιδιαζουσών τιμών πίνακα είναι στη συμπίεση εικόνων. Για παράδειγμα, αν ένας δορυφόρος μεταδίδει μια εικόνα που περιέχει 1000 x 1000 pixels και καθένα από αυτά γίνεται ψηφιακό, τότε η πληροφορία αυτή μπορεί να παρασταθεί με έναν πίνακα A τύπου 1000 x 1000. Η μετάδοση της πληροφορίας που δίνει ο πίνακας A αφορά 10^6 αριθμούς.

Αν χρησιμοποιήσουμε την ανάλυση SVD του πίνακα A με τη σχέση $A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_r u_r v_r^T$ πετυχαίνουμε μεγάλη μείωση του όγκου των δεδομένων. Στην πράξη, αν και το r είναι συνήθως αρκετά μεγάλο, πολλές από τις ιδιάζουσες τιμές είναι σχετικά μικρές ως προς τη μέγιστη ιδιάζουσα τιμή σ_1 , οπότε είναι δυνατό να τις αγνοήσουμε. Έτσι έχουμε μια σχετικά καλή προσέγγιση

$$A \approx \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \dots + \sigma_k u_k v_k^T$$

με k αρκετά μικρότερο από το r.

Π.χ. αν $k = 10$, τότε αρκεί να μεταδώσουμε μόνο τα 20 διανύσματα $\sigma_i u_i, v_i$, $i = 1, 2, \dots, 10$ τα οποία περιέχουν συνολικά 20000 αριθμούς.

Η διαδικασία που ακολουθείται είναι η εξής:

Διακριτοποιούμε την εικόνα χρησιμοποιώντας ορθογώνιο πλέγμα με πολλά και μικρά στοιχεία (pics). Σε κάθε στοιχείο, ανάλογα με τη σκιερότητα ή το χρωματισμό του, αντιστοιχίζεται ένας πραγματικός αριθμός. Η διάταξη των αριθμών σύμφωνα με την διάταξη των στοιχείων στην εικόνα, δημιουργεί ένα πολύ μεγάλο πίνακα A, με συνέπεια την πολύ δαπανηρή απεικόνισή του στον υπολογιστή. Η ανάλυση ιδιαζουσών τιμών του πίνακα A, δίνει τη δυνατότητα αναπαραγωγής της εικόνας χρησιμοποιώντας τον πίνακα A_k για αρκετά μικρό k χωρίς να υπάρχει μεγάλο σφάλμα στην εικόνα. Για την αποθήκευση απαιτούνται θέσεις μνήμης μόνο για τα ιδιάζοντα διανύσματα $u_1, \dots, u_k, v_1, \dots, v_k$ και για τις ιδιάζουσες τιμές $\sigma_1, \dots, \sigma_k$.

h) Ανάλυση Αντιστοιχιών (Correspondence Analysis – CA) των πινάκων συνάφειας:

Άλλη μια σημαντική εφαρμογή της ανάλυσης ιδιαζουσών τιμών πίνακα βρίσκεται στο χώρο της Στατιστικής.

Ο πίνακας συνάφειας $A \in \mathbb{R}^{n,m}$, $n > m$ είναι ο πίνακας συχνοτήτων ενός διδιάστατου δείγματος. Η Ανάλυση Αντιστοιχιών – CA διαμέσου της ανάλυσης ιδιαζουσών τιμών, αναλύει τον πίνακα A στο άθροισμα

$$A = \sigma_1 u_1 v_1^T + \sum_{i=2}^m \sigma_i u_i v_i^T .$$

Ο πρώτος όρος αντιστοιχεί στην ανεξαρτησία των μεταβλητών, ενώ ο δεύτερος μοντελοποιεί την απόκλιση αυτή.

Στόχος της CA είναι ο προσδιορισμός της μικρότερης δυνατής διάστασης που ερμηνεύει κατά τον καλύτερο τρόπο τη συσχέτιση των μεταβλητών. Δηλαδή, την προσέγγιση του $\sum_{i=2}^m \sigma_i u_i u_i^T$ από το $\sum_{i=2}^k \sigma_i u_i u_i^T$ όπου ο k είναι πρακτικά πολύ μικρότερος από τον m .

i) Αναπαράσταση εγγράφων κειμένου:

Η μέθοδος SVD δίνει το μαθηματικό υπόβαθρο για τη μέθοδο της Λανθάνουσας Σημασιολογικής Ανάλυσης (Latent Semantic Analysis) των εγγράφων, γνωστή στη διεθνή βιβλιογραφία ως LSA. Η μέθοδος προσπαθεί να επιλύσει τρία βασικά προβλήματα που χαρακτηρίζουν τα μοντέλα ανάκτησης. Τα προβλήματα αυτά είναι η συνωνυμία των όρων, η πολυσημία των όρων και η εξάρτηση μεταξύ τους.

Όπως αναφέραμε η LSA είναι μία μέθοδος ανάκτησης πληροφορίας βασισμένη στο εννοιολογικό περιεχόμενο των ερωτήσεων αναζήτησης.

Τα βασικά στάδια της LSA είναι τα εξής:

- Δημιουργία πίνακα όρων – εγγράφων
- Διάσπαση ιδιαζουσών τιμών (SVD)
- Μειωμένης διάστασης διάσπαση ιδιαζουσών τιμών (Reduced SVD)
- Αναπαράσταση – Συσχετισμός ερωτήσεων χρήστη

Εφαρμογές LSA:

- Πολυγλωσσικό λεξικό
- Μοντέλο ανθρώπινης σκέψης
- Αυτόματη βαθμολόγηση εκθέσεων
- Information filtering

Βιβλιογραφία

- [1] *Γραμμική Άλγεβρα και Αναλυτική Γεωμετρία*, Ανάργυρος Γ. Φελλούρης, Αθήνα, 2009
- [2] *Θέματα Ανάλυσης Πινάκων*, Παναγιώτης Ι. Ψαρράκος, Αθήνα, 2012
- [3] *Γραμμική Άλγεβρα Αναλυτική Γεωμετρία και Εφαρμογές*, Ν. Καδιανάκης – Σ. Καρανάσιος, Αθήνα, 2008
- [4] *Επιστημονικός Υπολογισμός Ι*, Ευστράτιος Γαλλόπουλος, Πάτρα, 2008
- [5] *Εγχειρίδιο Αριθμητικών Μεθόδων για Επιστημονικές Εφαρμογές με τη χρήση MATLAB και PYTHON*, Ηλίας Χούστης, Θεσσαλία, 2012
- [6] *Σημειώσεις Αριθμητικής Γραμμικής Άλγεβρας*, Β. Δουγαλής – Δ. Νούτσος – Α. Χατζηδήμος, Βόλος, 2007
- [7] *Ανάλυση Πίνακα σε Ιδιάζουσες Τιμές*, Wikipedia
- [8] *SVD Applications*, Stephen Boyd, Stanford, 2008
(http://www.columbia.edu/itc/applied/e3101/SVD_applications.pdf)
- [9] *Applications of the SVD*, Marc Spiegelman, New York, 2002
(<https://see.stanford.edu/materials/Isoldsee263/16-svd.pdf>)