



LINMA2380 Matrix Computations

SIMON DESMIDT

Academic year 2024-2025 - Q1



UCLouvain

Contents

1	Reminders	2
1.1	Algebraic structures	2
1.2	Matrix algebras	3
2	QR form	6
3	Unitary transformations and SVD	7
3.1	Introduction and definitions	7
3.2	Diagonalization by unitary transformations	7
3.3	Linear operator point of view	8
3.4	Polar decomposition - formal point of view	8
3.5	Projectors and generalized inverses - algebraic point of view	9
3.6	Least squares	9
3.7	Unitarily invariant matrix norms - geometric point of view	9
3.8	Canonical angles	9

Reminders

1.1 Algebraic structures

- A semigroupe is a set together with an associative binary operation $(E, +)$.
- A monoid is a semigroup with a neutral element.
- A group is a monoid in which every element has an inverse.
- A commutative group is a group whose binary operation is commutative.
- A ring is a triple $(E, +, \cdot)$ such that
 - $(E, +)$ is a commutative group;
 - (E, \cdot) is a monoid;
 - \cdot is distributive with respect to $+$.
- An integral domain is a commutative ring in which the product of any two nonzero elements is nonzero :

$$\forall x, y \in E, x, y \neq 0 \quad xy \neq 0$$

. This implies that the equation $ax = b$ with $a \neq 0$ has at most one solution.

- An Euclidean domain is an integral domain such that for every two elements in the domain, we can perform the Euclidean division:

$$\forall (a_1, a_2), \quad \exists (q, r) : \quad a_1 = a_2q + r \text{ with } r < a_2$$

- A field is a commutative ring $(E, +, \cdot)$ such that every $a \in E \setminus \{0\}$ has a multiplicative inverse.
- $(K, E, +)$ is a module over the ring $(K, +, \cdot)$ if
 - $(E, +)$ is a commutative group;
 - the external composition operation $\cdot : K \times E \rightarrow E$ satisfies
 - * $(a + b) \cdot x = a \cdot x + b \cdot x \quad a \cdot (x + y) = a \cdot x + a \cdot y$
 - * $a \cdot (b \cdot x) = (a \cdot b) \cdot x$
 - * $1 \cdot x = x$

- If, in addition to that, $(K, \cdot, +)$ is a field, then $(K, E, +)$ is a vector space over $(K, +, \cdot)$.
- $(K, E, +, \cdot)$ is an algebra if
 - $(K, E, +)$ is a module or a vector space;
 - the internal composition operation $\cdot : E \times E \rightarrow E$ is bilinear.

1.2 Matrix algebras

1.2.1 Product

Apart from the usual sum and product of two matrices, we can define the Hadamard and Kronecker products :

- Hadamard :

$$A_{m \times n} \odot B_{m \times n} := [a_{ij} \cdot b_{ij}]_{i,j=1}^{m,n}$$

- Kronecker :

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

A square matrix $A \in \mathbb{C}^{n \times n}$ is said normal if $AA^* = A^*$. In the real case, it is said to be orthogonal and $*$ is equivalent to the transpose. Furthermore, it is said to be unitary if it satisfies the relations $AA^* = I_n = A^*A$.

1.2.2 Determinant

We define the quasi-diagonals of a matrix as the n -tuples of elements of a matrix A , $a_{1j_1, 2j_2, \dots, nj_n}$ where the indices $\mathbf{j} = (j_1, \dots, j_n)$ constitute a permutation of the set $\{1, 2, \dots, n\}$. Thus a quasi-diagonal consists of n elements of the matrix A in such a way that no two of them lie in the same row or column of A . For each quasi-diagonal, we define the parity $t(\mathbf{j})$. It is the number of inversions $j_k > j_p$ for $k < p$ in \mathbf{j} .

- With the notation above, we define the determinant of a square matrix $A_{n \times n}$ as

$$\det(A) = \sum_{\mathbf{j}} (-1)^{t(\mathbf{j})} a_{1j_1} \cdot a_{2j_2} \cdot \dots \cdot a_{nj_n}$$

The determinant has the following properties :

- The determinant is multilinear in the rows of A :

$$\det \begin{bmatrix} a_{1:} \\ \vdots \\ b_{k:} + \lambda c_{k:} \\ \vdots \\ a_{n:} \end{bmatrix} = \det \begin{bmatrix} a_{1:} \\ \vdots \\ b_{k:} \\ \vdots \\ a_{n:} \end{bmatrix} + \det \begin{bmatrix} a_{1:} \\ \vdots \\ \lambda c_{k:} \\ \vdots \\ a_{n:} \end{bmatrix}$$

- The determinant is alternating in the rows of A : for $i \neq j$, $a_{i:} = a_{j:} \implies \det(A) = 0$
- $\det(I_n) = 1$, where I_n is the identity matrix.

$$\bullet \det \begin{bmatrix} a_{1:} \\ \vdots \\ a_{i:} \\ \vdots \\ a_{j:} \\ \vdots \\ a_{n:} \end{bmatrix} = - \det \begin{bmatrix} a_{1:} \\ \vdots \\ a_{j:} \\ \vdots \\ a_{i:} \\ \vdots \\ a_{n:} \end{bmatrix} \quad \bullet \det \begin{bmatrix} a_{1:} \\ \vdots \\ a_{i:} + \lambda a_{j:} \\ \vdots \\ a_{n:} \end{bmatrix} = \det \begin{bmatrix} a_{1:} \\ \vdots \\ a_{i:} \\ \vdots \\ a_{n:} \end{bmatrix}$$

- $\det(\lambda A) = \lambda^n \det(A)$
- for $i \neq j$, $a_{i:} = \lambda a_{j:} \implies \det(A) = 0$
- $\det(A^T) = \det(A)$
- $\det(A^*) = \overline{\det(A)}$, if $A \in \mathbb{C}^{n \times n}$

→ N.B.: any property of the determinant established for the rows of matrices also holds for the columns.

- The minor $A_{(kl)}$ of dimension $n - 1$ of a matrix $A_{n \times n}$ is the determinant of the submatrix obtained by removing the k th row and the l th column. From this, we can note the determinant as a linear combination of the elements of a row or column :

$$\det(A) = a_{1j}A_{1j}^c + a_{2j}A_{2j}^c + \cdots + a_{nj}A_{nj}^c \quad \det(A) = a_{i1}A_{i1}^c + a_{i2}A_{i2}^c + \cdots + a_{in}A_{in}^c$$

where the coefficient A_{kl}^c is called the cofactors of the corresponding element a_{kl} ¹

Laplace and Binet-Cauchy relations

For the pairs of p -tuples

$$\mathbf{i}_p := (i_1, \dots, i_p) \text{ and } \mathbf{j}_p := (j_1, \dots, j_p)$$

satisfying

$$1 \leq i_1 < \cdots < i_p \leq n \text{ and } 1 \leq j_1 < \cdots < j_p \leq n$$

we define the minors of order p of A as

$$A \begin{pmatrix} \mathbf{i}_p \\ \mathbf{j}_p \end{pmatrix} := \det[a_{i_k j_l}]_{k,l=1}^p \quad (1.1)$$

We also define the complementary cofactors of A as

$$A^c \begin{pmatrix} \mathbf{i}_p \\ \mathbf{j}_p \end{pmatrix} := (-1)^s A \begin{pmatrix} \mathbf{i}_p^c \\ \mathbf{j}_p^c \end{pmatrix} \quad (1.2)$$

¹ $A_{kl}^c = (-1)^{k+l} A_{(kl)}$.

where $s = \sum_{k=1}^p (i_k + j_k)$ and \mathbf{i}_p^c is the set complement of \mathbf{i}_p (same for \mathbf{j}_p).

Laplace Theorem:

Let A be a matrix of dimensions $n \times n$ and \mathbf{i}_p be a p -tuple of rows (and \mathbf{j}_p for the columns). Then, $\det(A)$ is equal to the sum of the products of all possible minors located in these rows/columns with their complementary cofactors:

$$\begin{cases} \det(A) = \sum_{\mathbf{j}_p} A \begin{pmatrix} \mathbf{i}_p \\ \mathbf{j}_p \end{pmatrix} A^c \begin{pmatrix} \mathbf{i}_p \\ \mathbf{j}_p \end{pmatrix} \\ \det(A) = \sum_{\mathbf{i}_p} A \begin{pmatrix} \mathbf{i}_p \\ \mathbf{j}_p \end{pmatrix} A^c \begin{pmatrix} \mathbf{i}_p \\ \mathbf{j}_p \end{pmatrix} \end{cases} \quad (1.3)$$

Binet-Cauchy Theorem:

Let \mathbf{m} be the m -tuple $(1, \dots, m)$. Let A and B be matrices of dimensions $m \times n$ and $n \times m$ respectively. If $m \leq n$, then

$$\det(AB) = \sum_{\mathbf{j}_m} A \begin{pmatrix} \mathbf{m} \\ \mathbf{j}_m \end{pmatrix} B \begin{pmatrix} \mathbf{j}_m \\ \mathbf{m} \end{pmatrix} \quad (1.4)$$

1.2.3 Inverse and rank

- The adjugate matrix of a square matrix $A_{n \times n}$ is defined as

$$\text{adj}(A) := [A_{ji}^c]_{i,j=1}^n$$

Then, for every square matrix $A_{n \times n}$, we have

$$A \cdot \text{adj}(A) = \det(A) I_n = \text{adj}(A) \cdot A \quad (1.5)$$

Every matrix $A_{m \times n}$ whose elements belong to a field \mathcal{F} can be brought to the following form by means of invertible (or elementary) transformations of rows and columns:

$$RAQ = \left(\begin{array}{c|c} I_r & 0_{r \times (n-r)} \\ \hline 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{array} \right) \quad (1.6)$$

The rank of a matrix $A_{m \times n}$ whose elements belong to a field \mathcal{F} is equal to the largest size of its nonzero minors. As a corollary, any non-singular matrix whose elements belong to a field \mathcal{F} can be written as a product of elementary transformations.

Schur complement:

Let $A_{n \times n}$ be an invertible submatrix of the matrix

$$M_{(n+p) \times (n+m)} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

Then the rank of M satisfies

$$\text{rank}(M) = n + \text{rank}(D - CA^{-1}B) \quad (1.7)$$

And the matrix $D - CA^{-1}B$ is called the Schur complement of M .

QR form

TODO

Unitary transformations and SVD

3.1 Introduction and definitions

- A unitary matrix is a matrix $U \in \mathbb{C}^{n \times n}$ such that $U^*U = I$, i.e. its column are orthogonal.
- An isometry is a matrix $U \in \mathbb{C}^{m \times n}$, $m \neq n$, such that $U^*U = I$. We have $\|Ux\| = \|x\|$.

3.2 Diagonalization by unitary transformations

The goal here is to have a matrix decomposition of the form

$$A = R \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q^{-1} \quad (3.1)$$

for any arbitrary matrix $A_{m \times n}$, and with R, Q being unitary (if A is complex) or orthogonal (if A is real). We limit ourselves here to transformation matrices that are isometries¹. This means that the invariants that we obtain characterize the way the matrix act on the norm of vectors.

Theorem 3.1. Every Hermitian² matrix $A \in \mathbb{C}^{n \times n}$ can be diagonalized by a unitary transformation $U \in \mathbb{C}^{n \times n}$:

$$U^*AU = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \\ \dots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \quad (3.2)$$

with $\lambda_i \in \mathbb{R}$.

Theorem 3.2. The eigenvalues of a Hermitian matrix $A \in \mathbb{C}^{n \times n}$ are invariant under unitary similarity transformations:

$$B = U^*AU \quad (3.3)$$

Every class of equivalence defined by this transformation group has a unique canonical representative which is the diagonal matrix Λ with the eigenvalues of A decreasing along the diagonal.

¹To define.

² $A = A^*$

Theorem 3.3 (Singular Value Decomposition). For every matrix $A \in \mathbb{C}^{m \times n}$, there exist unitary transformations $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ such that

$$A = U\Sigma V^* \quad \Sigma = \left(\begin{array}{ccc|ccc} \sigma_1 & & & & & \\ & \ddots & & & & \\ 0 & & \sigma_r & & & \\ \hline & & & 0_{(m-r) \times r} & & \\ & & & & 0_{(m-r) \times (n-r)} & \end{array} \right) \quad (3.4)$$

with real positive singular values $\sigma_1 \geq \dots \geq \sigma_r > 0$. The value r and the r -tuple $(\sigma_1, \dots, \sigma_r)$ are uniquely defined and, as a consequence, the matrix Σ constitutes a canonical form under unitary transformations, i.e. under transformations of the forme $B = \tilde{U}^* A \tilde{V}$. Where \tilde{U}, \tilde{V} are two unitary matrices.

Properties:

- If the matrix A is real, U, V are orthogonal matrices;
- The transformations U, V diagonalize the matrices AA^* and A^*A respectively, since $U^*AA^*A = \Sigma\Sigma^T$, $V^*A^*AV = \Sigma^T\Sigma$, and the columns of U, V are the eigenvectors of AA^* and A^*A respectively.
- The transformations U, V are not uniquely defined.

3.3 Linear operator point of view

We define the compact SVD form: $A = U_1 \Sigma_r V_1^*$, to have Σ_r invertible. In this form, Σ_r is $r \times r$, r being the number of nonzero singular values. U_1 contains the r first columns of U and V_1^* the r first lines of V^* . The other columns (resp. rows) of U (resp. V) are denoted by the matrix U_2 (resp. V_2^*).

Definition 3.4. If $\mathcal{X}_1, \mathcal{X}_2$ are subspaces of \mathbb{R}^n such that their intersection is the origin, then we note $\mathcal{X}_1 \oplus \mathcal{X}_2 = \mathcal{X}_1 + \mathcal{X}_2$ the direct sum of the two spaces.

Any vector $x \in \mathcal{X}_1 \oplus \mathcal{X}_2$ has a unique decomposition $x = x_1 + x_2$, $x_i \in \mathcal{X}_i$.
For the SVD, we have

$$\mathcal{X}_1 = \text{Im}(V_1) \quad \mathcal{X}_2 = \text{Im}(V_2) = \text{Ker}(A) \quad (3.5)$$

$$\mathcal{Y}_1 = \text{Im}(U_1) = \text{Im}(A) \quad \mathcal{Y}_2 = \text{Im}(U_2) \quad (3.6)$$

3.4 Polar decomposition - formal point of view

Any matrix $A_{n \times n}$ can be expressed in the following form:

$$A = \underbrace{U\Sigma U^*}_{=:H_1} UV^* = H_1 Q = H_1 \exp(iH_2) \quad (3.7)$$

with H_1 a positive definite Hermitian matrix, Q unitary and H_2 also Hermitian.

3.5 Projectors and generalized inverses - algebraic point of view

Definition 3.5. A projector is a matrix $P \in \mathbb{C}^{n \times n}$ such that $P^2 = P$. It is said to be orthogonal if $\forall x, (Px)^*(x - Px) = 0$.

Theorem 3.6. Any projector P can be written $P = XY^*$ with $Y^*X = I_r$, r being the rank of P . If P is orthogonal, then $X = Y$.

- $Im(P) = Ker(P^\perp)$
- $P = P^*$

3.6 Least squares

Theorem 3.7. Given a linear system $Ax = y$, the generalized inverse $A^I = V_1 \Sigma_r^{-1} U_1^*$ gives $x^* = A^I y$ the solution minimizing the norm of $Ax - y$. If there are more than one such solution, it returns the one of smallest norm.

3.7 Unitarily invariant matrix norms - geometric point of view

A matrix norm is unitarily invariant if, for every $A \in \mathbb{C}^{m \times n}$, we have $\|A\| = \|U^*AV\|$ if U, V are unitary.

The 2-norm and the Frobenius norm of $A \in \mathbb{C}^{m \times n}$ are unitarily invariant.

$$\|A\|_2 := \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} \quad \|A\|_F := \left(\sum_{i,j} |a_{i,j}|^2 \right)^{1/2} \quad (3.8)$$

3.8 Canonical angles

Theorem 3.8. Given two subspaces $\mathcal{S}_i \subseteq \mathbb{C}^n$ ($i = 1, 2$), there exist orthonormal bases given by the columns of \hat{S}_i respectively, and satisfying

$$\hat{S}_1^* \hat{S}_2 = \left(\begin{array}{ccc|ccc} \sigma_1 & & & 0 & & \\ & \ddots & & & & \\ 0 & & & \sigma_r & & \\ \hline & & 0_{(r_1-r) \times r} & & 0_{(r_1-r) \times (r_2-r)} & \end{array} \right) \quad 1 \geq \sigma_1 \geq \dots \geq \sigma_r > 0 \quad (3.9)$$

Add the paper sheet of notes.