LINMA2380 Matrix Computations

Simon Desmidt

Contents

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 distirbutive with respect to $+$.

- An integral domain is a commutative ring in which the product of any two nonzero elements in nonzero :

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

  . 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:

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

- 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 $\cdot : K \times E \to E$ satisfies
    * $(a + b) \cdot x = a \cdot x + b \cdot x \qquad a \cdot (x + y) = a \cdot x + a \cdot y$
    * $a \cdot (b \cdot x) = (a \cdot b) \cdot x$
    * $1 \cdot x = x$

- 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 $\cdot : E \times E \to E$ 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.

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

- $\det(\lambda A) = \lambda^n \det(A)$

- for $i \neq j$, $a_{i:} = \lambda a_{j:} \implies \det(A) = 0$

- $\det(A^T) = \det(A)$

- $\det(A^*) = \overline{\det(A)}$, if $A \in \mathbb{C}^{n \times n}$

$\rightarrow$ 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 $k$the row and the lth 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

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

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

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

We also define the complementary cofactors of A as

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

---

[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 ne 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 possibles minors located in these rows/columns with their complementary cofactors:

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

Binet-Cauchy Theorem:

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

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

1.2.3 Inverse and rank

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

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

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

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

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:

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

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

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

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

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

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 isometries[1]. This means that the invariants that we obtain characterize the way the matrix act on the norm of vectors.

Theorem 3.1. Every Hermitian[2] matrix $A \in \mathbb{C}^{n \times n}$ can be diagonalized by a unitary transformation $U \in \mathbb{C}^{n \times n}$:

$$U^*AU = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \dots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \tag{3.2}$$

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:

$$B = U^*AU \tag{3.3}$$

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.

---

[1] To 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

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

with real positive singular values $\sigma_1 \geq \cdots \geq \sigma_r > 0$. The value r and the r-tuple $(\sigma_1, \ldots, \sigma_r)$ 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 $U^* A A^* A = \Sigma \Sigma^T$, $V^* A^* A V = \Sigma^T \Sigma$, 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

$$\mathcal{X}_1 = Im(V_1) \qquad\qquad \mathcal{X}_2 = Im(V_2) = Ker(A) \tag{3.5}$$
$$\mathcal{Y}_1 = Im(U_1) = Im(A) \qquad\qquad \mathcal{Y}_2 = Im(U_2) \tag{3.6}$$

3.4 Polar decomposition - formal point of view

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

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

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.

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

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

$$\hat{S}_1^* \hat{S}_2 = \begin{pmatrix} \sigma_1 & & 0 & \\ & \ddots & & 0_{r \times (r_2 - r)} \\ 0 & & \sigma_r & \\ \hline & 0_{(r_1 - r) \times r} & & 0_{(r_1 - r) \times r_2 - r} \end{pmatrix} \qquad 1 \geq \sigma_1 \geq \cdots \geq \sigma_r > 0 \tag{3.9}$$

Add the paper sheet of notes.

3.9 Variational problems

Theorem 3.9. For a Hermitian matrix $H \in \mathbb{C}^{n \times n}$, the Rayleigh quotient is defined as

$$R(x) := \frac{\langle Hx, x \rangle}{\langle x, x \rangle} = \frac{x^* H x}{x^* x} \qquad x \neq 0 \in \mathbb{C}^n \tag{3.10}$$

The Rayleigh quotient of a Hermitian matrix $H \in \mathbb{C}^{n \times n}$ is real and satisfies

$$\lambda_{\min}(H) \leq R(x) \leq \lambda_{\max}(H) \tag{3.11}$$

Furthermore, supposing that $\lambda_1 \geq \cdots \geq \lambda_n$, we have

$$\lambda_n = \min_{x \neq 0} R(x) \qquad \lambda_1 = \max_{x \neq 0} R(x) \tag{3.12}$$

Lemma 3.10. Let $\mathcal{S}_j \subseteq \mathbb{C}^n$ be a subspace of dimension j. Then, it holds that

$$\min_{x \neq 0 \in \mathcal{S}_j} R(x) \leq \lambda_j \qquad \max_{x \neq 0 \in S_j} R(x) \geq \lambda_{n-j+1} \tag{3.13}$$

Theorem 3.11 (Courant-Fisher)**.** For any Hermitian matrix $H \in \mathbb{C}^{n \times n}$, the Rayleigh quotient $R(x)$ satisfies

$$\lambda_j = \max_{\mathcal{S}_j} \min_{x \neq 0 \in \mathcal{S}_j} R(x) \qquad \lambda_{n-j+1} = \min_{\mathcal{S}_j} \max_{x \neq 0 \in \mathcal{S}_j} R(x) \tag{3.14}$$

Theorem 3.12. The singular values of an arbitrary matrix $A \in \mathbb{C}^{m \times n}$ are given by

$$\sigma_j(A) = \max_{\mathcal{S}_j} \min_{x \neq 0 \in \mathcal{S}_j} \frac{\|Ax\|_2}{\|x\|_2} \tag{3.15}$$

$$\sigma_{n-j+1}(A) = \min_{\mathcal{S}_j} \max_{x \neq 0 \in \mathcal{S}_j} \frac{\|Ax\|_2}{\|x\|_2} \tag{3.16}$$

The following theorem is a major application of the SVD, as it allows to store a matrix with much less information that it contains.

Theorem 3.13. Let $A \in \mathbb{C}^{m \times n}$ be a matrix of rank r. The best approximation of A by a matrix $B \in \mathbb{C}^{m \times n}$ of rank $s < r$ satisfies

$$\min_{\text{rank}(B) \leq s} \|A - B\|_2 = \sigma_{s+1}(A) \tag{3.17}$$

Theorem 3.14 (Eckart-Young)**.** Furthermore, the matrix A from the last theorem also satisfies

$$\min_{\text{rank}(B) \leq s} \|A - B\|_F^2 = \sigma_{s+1}^2 + \cdots + \sigma_r^2 \tag{3.18}$$

Eigenvalues, eigenvectors and similarity transformations

The eigenvalues of a matrix are invariant under similarity transformations. The similiary transformations $A \to TAT^{-1}$ define an equivalence class of matrices and every matrix A_T belonging to the similarity class of A has the same eigenvalues.

Theorem 4.1 (Schur). Every matrix $A \in \mathbb{C}^{n \times n}$ can be upper triangularized under unitary similarity transformations:

$$U^* A U = \begin{bmatrix} \lambda_1 & \times & \cdots & \times \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \times \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix} =: A_S \tag{4.1}$$

where the diagonal of A_S consists of the eigenvalues of A.

- If A is Hermitian, then A_S is Hermitian, and thus diagonal and real. This is its canonical form.

- The eigenvalues in the Schur form can be ordered.

- If $A \in \mathbb{R}^{n \times n}$, the eigenvalues and eigenvectors can be complex. Then, their complex conjugates are also eigenvalues and eigenvectors of A.

Definition 4.2. A normal matrix is a square matrix $A \in \mathbb{C}^{n \times n}$ satisfying $AA^* = A^*A$.

Theorem 4.3. A matrix $A \in \mathbb{C}^{n \times n}$ is normal if and only if it is diagonalizable under unitary similarity transformations: $A = U \Lambda U^*$.

4.1 Invariant subspaces

Definition 4.4. A subset $\mathcal{X} \subseteq \mathbb{C}^n$ is an invariant subspace under the operator $A \in \mathbb{C}^{n \times n}$ if $A\mathcal{X} \subseteq \mathcal{X}$.

Theorem 4.5. Let $\mathcal{X} \subseteq \mathbb{C}^n$ be a subspace of dimension k. Let $X \in \mathbb{C}^{n \times k}$ be such that the columns of X form a basis of $\mathcal{X}$, and let X_c be a completion of X such that $T := [X|X_c]$ is non-singular. Then, the following three propositions are equivalent:

- $A\mathcal{X} \subseteq \mathcal{X}$;

- $AX = XA_{11}$;

- $T^{-1}AT = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0_{(n-k)\times k} & A_{22} \end{array}\right]$

where $A_{11} \in \mathbb{C}^{k\times k}$, $A_{12} \in \mathbb{C}^{k\times(n-k)}$ and $A_{22} \in \mathbb{C}^{(n-k)\times(n-k)}$.

Theorem 4.6 (Real Schur form). Every matrix $A \in \mathbb{R}^{n\times n}$ can be almost triangularized under real similarity transformations $U \in \mathbb{R}^{n\times n}$, and with blocks of dimension 1×1 or 2×2:

$$U^T A U = \begin{bmatrix} A_{11} & \times & \dots & \times \\ & A_{22} & \ddots & \vdots \\ & & \ddots & \times \\ & & & A_{kk} \end{bmatrix} \qquad A_{ii} \in \mathbb{R}^{1\times 1} \cup \mathbb{R}^{2\times 2} \tag{4.2}$$

4.2 Jordan canonical form

Theorem 4.7. Every matrix $A \in \mathbb{C}^{n\times n}$ admits a block-diagonal form under similarity transformations:

$$T^{-1}AT = A_S = diag\{A_{11}, \dots, A_{kk}\} \tag{4.3}$$

where each block A_{ii} has only one eigenvalue (with multiplicity possibly larger than 1).

Theorem 4.8. Every matrix $A \in \mathbb{C}^{n\times n}$ can be transformed by similarity transformations into a block-diagonal form:

$$T^{-1}AT = diag\{J_1(\lambda), \dots, J_k(\lambda)\} \tag{4.4}$$

where each $J_i(\lambda) \in \mathbb{C}^{n_i\times n_i}$ is a Jordan block:

$$J_i(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix} \tag{4.5}$$

Corollary 4.9. Two matrices $A, B \in \mathbb{C}^{n\times n}$ are similar iff they have the same Jordan form.

For a real matrix $A \in \mathbb{R}^{n\times n}$, we can separate the real and imaginary parts of the eigenvalues: $\lambda = \alpha + \beta j$. Separating the eigenvectors $v = x + iy$:

$$A \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \tag{4.6}$$

Corollary 4.10. Any real matrix $A \in \mathbb{R}^{n \times n}$ has an invariant space of dimensionn 1 or 2:

$$
\begin{bmatrix}
\alpha & \beta & 1 & 0 & & & & \\
-\beta & \alpha & 0 & 1 & & & & \\
& & \alpha & \beta & 1 & 0 & & \\
& & -\beta & \alpha & 0 & 1 & & \\
& & & & \ddots & & \ddots & \\
& & & & & 1 & 0 \\
& & & & & 0 & 1 \\
& & & & & \alpha & \beta \\
& & & & & -\beta & \alpha
\end{bmatrix}
\tag{4.7}
$$

Definition 4.11. Given a polynomial $p(x) = p_0 + p_1 x + \cdots + p_d x^d$, we define the matrix polynomial $p(A) = p_0 I + p_1 A + \cdots + p_d A^d$.

Theorem 4.12. If $A = TJT^{-1}$, $p(A) = Tp(J)T^{-1}$, and thus

$$
\lambda_i(p(A)) = p(\lambda_i(A))
\tag{4.8}
$$

Corollary 4.13. The characteristic polynomial of A satisfies $\chi(A) = \prod_i (A - \lambda_i I)^{m_i} = 0$, with m_i the algebraic multiplicity of the eigenvalue λ_i.

Theorem 4.14. (Bendixson) The eigenvalues $\alpha + \beta j$ of a matrix $A \in \mathbb{C}^{n \times n}$ are in the rectangle

$$
\lambda_{\min}\left(\frac{A + A^*}{2}\right) \leq \alpha \leq \lambda_{\max}\left(\frac{A + A^*}{2}\right)
\tag{4.9}
$$

$$
\lambda_{\min}\left(\frac{A - A^*}{2j}\right) \leq \beta \leq \lambda_{\max}\left(\frac{A - A^*}{2j}\right)
\tag{4.10}
$$

Theorem 4.15. Let $A \in \mathbb{C}^{n \times n}$. Its eigenvalues are all in the union of the Gershgorin circles:

$$
|z - a_{pp}| \leq \sum_{k \neq p} |a_{pk}|
\tag{4.11}
$$

Polynomial matrices

5.1 Group of elementary matrices

1. Permutations: P such that $P^T = P^{-1}$;

2. Eliminations: $(I + p(x)E_{ij})^{-1} = I - p(x)E_{ij}$ where E_{ij} is such that $e_{ij} = 1$ and all other elements are 0;

3. Scaling matrices : The identity with a polynomial on one of the term of the diagonal.

Type I, II, and III with a constant polynomial are unimodular matrices.

5.2 Smith form

Definition 5.1. A polynomial matrix $E(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ is unimodular if its determinant is a nonzero constant.

From this it comes that the invertible polynomial matrices are exactly the unimodular matrices.

Algorithm 1 Euclide's algorithm

1: Input: $a \geq b \in \mathbb{Z}$.
2: Output: GCD(a,b).
3: $a = bq_1 + r_1$
4: $i = 1$;
5: $r_0 = b$;
6: Repeat if $r_i \neq 0$;
7: If $r_i = 0$, $r_{i+1} := r_{i-1} - r_i a_{i+1}$.

Lemma 5.2. The Euclide algorithm terminates in finite time T, with $r_T = 0$ and $r_{T-1} = GCD(a, b)$.

Theorem 5.3. Euclid-Stevin For every two polynomials $a(\lambda), b(\lambda) \in \mathbb{C}[\lambda]$, there exists a unimodular transformation $U(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ such that

$$\begin{bmatrix} a(\lambda) \\ b(\lambda) \end{bmatrix} = U(\lambda) \begin{bmatrix} d(\lambda) \\ 0 \end{bmatrix} \tag{5.1}$$

where $d(\lambda)$ is the greatest common divisor (gcd) of $a(\lambda)$ and $b(\lambda)$.

Corollary 5.4. For every n polynomials $p_1(\lambda), \ldots, p_n(\lambda) \in \mathbb{C}[\lambda]$, there exists a unimodular transformation $Q(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ such that

$$Q(\lambda) \begin{bmatrix} p_1(\lambda) \\ \vdots \\ p_n(\lambda) \end{bmatrix} = \begin{bmatrix} d(\lambda) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5.2}$$

where $d(\lambda)$ is the gcd of $p_1(\lambda), \ldots, p_n(\lambda)$.

Theorem 5.5. Hermite Every polynomial matrix $P(\lambda) \in \mathbb{C}^{m \times n}[\lambda]$ can be transformed into a quasi-triangular form

$$M(\lambda)P(\lambda)N = \left(\begin{array}{ccc|ccc} p_1(\lambda) & \dots & \times & \times & \dots & \times \\ & \ddots & \vdots & \vdots & & \vdots \\ & & p_r(\lambda) & \times & \dots & \times \\ \hline & 0_{(m-r)\times r} & & & 0_{(m-r)\times(n-r)} & \end{array} \right) \tag{5.3}$$

where $M(\lambda) \in \mathbb{C}^{m \times m}[\lambda]$ is unimodular and $N \in \mathbb{R}^{n \times n}$ is a permutation matrix.

Proposition 5.6. Let $P \in \mathbb{Z}^{n \times n}$, M, N unimodular and $a \in \mathbb{Z}$. All $i \times i$ minors of P are divisible by a iff all minors of MPN are divisible by a.

Theorem 5.7. Smith Every polynomial matrix $P(\lambda) \in \mathbb{C}^{m \times n}[\lambda]$ can be reduced by unimodular transformations $M(\lambda) \in \mathbb{C}^{m \times m}[\lambda]$ and $N(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ to a quasi-diagonal form

$$M(\lambda)P(\lambda)N(\lambda) = \left(\begin{array}{ccc|c} e_1(\lambda) & \dots & \times & \\ & \ddots & \vdots & 0_{r \times (n-r)} \\ & & e_r(\lambda) & \\ \hline & 0_{(m-r)\times r} & & 0_{(m-r)\times(n-r)} \end{array} \right) \tag{5.4}$$

where $e_i(\lambda)$ divides $e_{i+1}(\lambda)$, which we denote $e_i(\lambda)|e_{i+1}(\lambda)$ $(i = 1, \ldots, r-1)$.

Theorem 5.8. Let $i \leq k \leq n$.

$$\prod_{1 \leq j \leq k} e_j = GCD \text{ of all minors of size } k \tag{5.5}$$

This means that the Smith form is canonical.

Definition 5.9. The normal rank of a polynomial matrix $P(\lambda) \in \mathbb{C}^{m \times n}[\lambda]$ is the order of its largest nonzero minor.

→ Note: the Smith form is a canonical form under the group of unimodular transformations on the left and on the right. These transformations are equivalences, and two polynomial matrices are thus equivalent iff they have the same Smith form.

5.2.1 Smith form of a Jordan block

Lemma 5.10. The Smith form of $\lambda I - A$ is the Smith form of $\lambda I - J_A$, J_A being the Jordan form of A.

Proposition 5.11. The Smith form of a Jordan block of size n

$$
\begin{pmatrix}
\lambda - \lambda_i & 1 & & & \\
& \lambda - \lambda_i & 1 & & \\
& & \ddots & \ddots & \\
& & & \lambda - \lambda_i & 1 \\
& & & & \lambda - \lambda_i
\end{pmatrix}
$$

is

$$
\begin{pmatrix}
1 & & & \\
& 1 & & \\
& & \ddots & \\
& & & (\lambda - \lambda_i)^n
\end{pmatrix}
\tag{5.6}
$$

For a Jordan matrix with several blocks of different sizes n_j with the same eigenvalue, the Smith form is

$$
\begin{pmatrix}
1 & & & & & & \\
& 1 & & & & & \\
& & \ddots & & & & \\
& & & (\lambda - \lambda_i)^{n_1} & & & \\
& & & & (\lambda - \lambda_i)^{n_2} & & \\
& & & & & \ddots & \\
& & & & & & (\lambda - \lambda_i)^{n_k}
\end{pmatrix}
\tag{5.7}
$$

Theorem 5.12. Let $A \in \mathbb{C}^{n \times n}$. The Smith form of $\lambda I - A$ is $diag(e_1(\lambda), e_2(\lambda), \ldots, e_n(\lambda))$, where $e_{n-k+1}(\lambda)$ contains a factor $(\lambda - \lambda_i)^d$. if there are more than k Jordan blocks of eigenvalue λ_i and d is the size of the kth largest Jordan block.

Nonnegative matrices

Definition 6.1. A nonnegative matrix A is such that all its entries are nonnegative.

Definition 6.2. A nonnegative matrix $A \in \mathbb{R}_+^{n \times n}$ is irreducible if $(I + A)^{n+1} > 0$.

Definition 6.3. A nonnegative matrix $A \in \mathbb{R}_+^{n \times n}$ is primitive if there exists $m \in \mathbb{Z}_+$ such that $A^m > 0$.

Theorem 6.4. In the case of A irreducible, this theorem is referred as the Perron-Frobenius theorem.

Theorem 7.3: Spectral properties of nonnegative matrices

Let $A \in \mathbb{R}^{n \times n}$, and let $\mathrm{diag}(J_1(\lambda_1), \ldots, J_k(\lambda_k))$ be its Jordan canonical form, where for each $i \in \{1, \ldots, k\}$, $J_i(\lambda_i)$ is a $n_i \times n_i$ Jordan block.

- If $A \geq 0$, then there is $i \in \{1, \ldots, k\}$ such that

    (a) $\rho(A) = \lambda_i$;

    (b) for all $j \in \{1, \ldots, k\} \setminus \{i\}$, $|\lambda_j| < |\lambda_i|$ or $n_j \leq n_i$.

    (c) there is $\mathbf{x} \geq 0 \in \mathbb{R}^n$ that is an eigenvector for λ_i, i.e., $\mathbf{x} \neq 0$ and $A\mathbf{x} = \lambda_i \mathbf{x}$.

- If $A \geq 0$ is irreducible, then $\rho(A) > 0$ and there is $i \in \{1, \ldots, k\}$ such that (a)–(b) hold and

    (d) $n_i = 1$;

    (e) for all $j \in \{1, \ldots, k\} \setminus \{i\}$, $\lambda_j \neq \lambda_i$;

    (f) there is $\mathbf{x} > 0 \in \mathbb{R}^n$ that is an eigenvector for λ_i, i.e., $\mathbf{x} \neq 0$ and $A\mathbf{x} = \lambda_i \mathbf{x}$;

    (g) any eigenvector $\mathbf{x}$ associated to an eigenvalue $\lambda_j \neq \rho(A)$ satisfies $\mathbf{x} \not\geq 0$.

- If $A \geq 0$ is primitive, then $\rho(A) > 0$ and there is $i \in \{1, \ldots, k\}$ such that (a), (d)–(g) hold and

    (h) for all $j \in \{1, \ldots, k\} \setminus \{i\}$, $|\lambda_j| < |\lambda_i|$.

If $A \geq 0$ is irreducible, then (h) is sufficient for A to be primitive.

Theorem 6.5. Ax is monotonous in x if $x \geq x' \implies Ax \geq Ax'$. If, for some x, $Ax \geq Mx$, then $A^t x \geq M^t x$ and $|\lambda_1(A)| \geq M$.

Proposition 6.6. If $A^k > 0$ for some k, then the optimal x^* in

$$\max_{\|x\|=1} \{r : Ax \geq rx\} \tag{6.1}$$

is such that $Ax^* = r^* x^*$.

Corollary 6.7. Let $A \geq 0 \in \mathbb{R}^{n \times n}$ be primitive and assume that $\rho(A) = 1$. Let $x > 0 \in \mathbb{R}^n$ be right and left eigenvectors of A associated to the eigenvalue 1 satisfying $y^T x = 1$. Then, $\lim_{m \to \infty} A^m = xy^T$.

Theorem 6.8. If $A \geq 0 \in \mathbb{R}^{n \times n}$ is irreducible and stochastic, then $\rho(A) = 1$ and 1 is a simple eigenvalue of A. Hence, there is a unique vector $x > 0 \in \mathbb{R}^n$ such that $Ax = x$ and $u^T x = 1$.

Corollary 6.9. If $A \geq 0 \in \mathbb{R}^{n \times n}$ is primitive and stochastic, then $\lim_{m \times \infty} A^m = xu^T$, where x is as in Theorem 6.8.

6.1 Graph-theoretic characterization

Definition 6.10. Given $A \geq 0 \in \mathbb{R}^{n \times n}$, the graph G_A is strongly connected if for any pair of nodes $(i, j) \in \{1, \ldots, n\}^2$, there is a path from i to j in G_A.

Definition 6.11. Given $A \geq 0 \in \mathbb{R}^{n \times n}$, the graph G_A is aperiodic if the gcd of all cycles in G_A is 1.
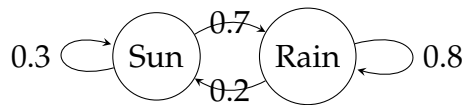
Lemma 6.12. Let $A \geq 0 \in \mathbb{R}^{n \times n}$, $(i, j) \in \{1, \ldots, n\}^2$ and $m \in \mathbb{Z}_+$. It holds that $[A^m]_{ij} > 0$ iff there exists a path from i to j of length m in G_A.

Theorem 6.13. A matrix $A \geq 0 \in \mathbb{R}^{n \times n}$ is irreducible iff G_A is strongly connected.

Theorem 6.14. A matrix $A \geq 0 \in \mathbb{R}^{n \times n}$ is primitive iff G_A is strongly connected and aperiodic.

Semigroups of matrices

7.1 Markov chains

The state of the Markov chain is $x_t = \begin{pmatrix} S_t := \text{ probability that there is sun at day } t \\ R_t := 1 - S_t \end{pmatrix}$.

With this example, we have the recurrence

$$x_{t+1} = \begin{pmatrix} 0.3 & 0.2 \\ 0.7 & 0.8 \end{pmatrix} x_t \tag{7.1}$$

$\rightarrow$ Note: $|\lambda_1(A)| = 1$ to keep $S_t + R_t = 1 \; \forall t$.

As t increases, x tends to a constant vector $(1/2, 1/2)^T$ whatever the initial conditions.

7.2 Definitions

Definition 7.1. The joint spectral radius of a bounded set of matrices $\Sigma \subseteq \mathbb{R}^{n \times n}$ is defined by

$$\rho(\Sigma) = \limsup_{t \to \infty} \rho_t(\Sigma) = \lim_{t \to \infty} \hat{\rho}_t(\Sigma) \tag{7.2}$$

with

$$\hat{\rho}_t(\Sigma) := \sup\{\|A\|^{1/t} : A \in \Sigma^t\} \tag{7.3}$$

Definition 7.2. Using

$$\begin{cases} \check{\rho}_t(\Sigma) := \inf\{\|A\|^{1/t} : A \in \Sigma^t\} \\ \rho_t(\Sigma) := \inf\{\rho(A)^{1/t} : A \in \Sigma^t\} \end{cases} \tag{7.4}$$

The joint spectral subradius of a set of matrices $\Sigma \subseteq \mathbb{R}^{n \times n}$ is defined by

$$\check{\rho}(\Sigma) = \lim_{t \to \infty} \check{\rho}_t(\Sigma) = \lim_{t \to \infty} \rho_t \tag{7.5}$$

7.3 Theorems

Theorem 7.3. For any bounded set of matrices $\Sigma \subseteq \mathbb{R}^{n \times n}$, the values $\check{\rho}(\Sigma)$ and $\rho(\Sigma)$ are equal.

Theorem 7.4. For any bounded set of matrices $\Sigma \subseteq \mathbb{R}^{n \times n}$, there exists a left-infinite product $\ldots A_2 A_1$ that does not converge to zero iff $\rho(\Sigma) \geq 1$.

Corollary 7.5. For any bounded set of matrices $\Sigma \subseteq \mathbb{R}^{n \times n}$, the corresponding switched dyanmical system like the one in the Markov chains is stable iff $\rho(\Sigma) < 1$.

Corollary 7.6. For any bounded set of matrices $\Sigma \subseteq \mathbb{R}^{n \times n}$, there is an infinite product of these matrices reaching the joint spectral radius, i.e. there is a sequence of matrices $A_0, A_1, \ldots$ in Σ such that

$$\lim_{t \to \infty} \|A_t \ldots A_1 A_0\|^{1/t} = \rho(\Sigma) \tag{7.6}$$

Corollary 7.7. For any set of matrices $\Sigma \subseteq \mathbb{R}^{n \times n}$, there is an infinite product of these matrices reaching the joint spectral subradius, i.e. there is a sequence of matrices $A_0, A_1, \ldots$ in Σ such that

$$\lim_{t \to \infty} \|A_t \ldots A_1 A_0\|^{1/t} = \check{\rho}(\Sigma) \tag{7.7}$$

Summary

Relation	Definition set	Definition $A \sim B$	Invariant $A \sim B$	Canonical element
Equivalence relation under elementary transformations on the left and on the right	$\mathbb{C}^{m \times n}$	$\exists P_1 \in K^{m \times m}, P_2 \in \mathbb{C}^{n \times n}$ inversible : $B = P_1 A P_2$	Rank $\mathrm{rank}(A) = \mathrm{rank}(B)$	$\left(\begin{array}{c\|c} I_r & 0_{r \times (n-r)} \\ \hline 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{array} \right)$
Similarity relation	$\mathbb{C}^{n \times n}$	$\exists P \in \mathbb{C}^{n \times n}$ inversible : $B = P^{-1} A P$	Jordan canonical form $J(A) = J(B)$	$\begin{bmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_p}(\lambda_p) \end{bmatrix}$
Congruence relation	$\mathbb{R}^{n \times n}$ symmetric	$\exists P \in \mathbb{R}^{n \times n}$ inversible : $B = P^{\top} A P$	Inertia $\mathrm{Inert}(A) = \mathrm{Inert}(B)$	$\begin{bmatrix} I_{\mathrm{ind}_+(A)} & & \\ & -I_{\mathrm{ind}_-(A)} & \\ & & 0_{\mathrm{ind}_0(A)} \end{bmatrix}$
Equivalence relation under unitary transformations on the left and on the right	$\mathbb{C}^{m \times n}$	$\exists U \in \mathbb{C}^{m \times m}, V \in \mathbb{C}^{n \times n}$ unitary: $B = U A V^*$	Singular values	$\left(\begin{array}{cc\|c} \sigma_1 & 0 & \\ & \ddots & 0_{r \times (n-r)} \\ 0 & \sigma_r & \\ \hline 0_{(m-r) \times r} & & 0_{(m-r) \times (n-r)} \end{array} \right)$
Equivalence relation under unimodular transformations on the left and on the right	$\mathbb{C}^{m \times n}[\lambda]$	$\exists M(\lambda) \in \mathbb{C}^{m \times m}[\lambda], N(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ unimodular: $B(\lambda) = M(\lambda) A(\lambda) N(\lambda)$	Smith normal form	$\left(\begin{array}{cc\|c} e_1(\lambda) & 0 & \\ & \ddots & 0_{r \times (n-r)} \\ 0 & e_r(\lambda) & \\ \hline 0_{(m-r) \times r} & & 0_{(m-r) \times (n-r)} \end{array} \right)$ where $e_i(\lambda)$ divides $e_{i+1}(\lambda)$ $(i = 1, \ldots, r-1)$.