

State-Space Control Fundamentals (00860721)

lecture no. 1

Leonid Mirkin

The Stephen B. Klein Faculty of Aerospace Engineering
Technion—IIT



Course info

Course web site: <http://leo.technion.ac.il/Courses/SSCF/>

Prerequisite:

- Control Theory (00840738)
- some proficiency in Linear Algebra and Dynamic Systems is expected

Grading policy:

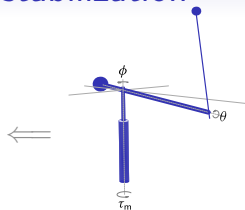
- quizzes during lectures (average of 5 best out of 6, *magen*): 20%
(provided the final exam grade is at least 55)
- homework (average of 3 best out of 4, *tokef*): 40%
- Final exam: 40–60%

The use of books / lecture notes **not** permitted during exams.

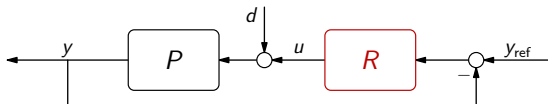
Example 1: inverted Furuta pendulum stabilization

plant $P : \tau_m \mapsto \phi$ with

$$P(s) = \frac{97.637(s + 5.382)(s - 5.27)}{s(s + 0.7318)(s + 6.351)(s - 5.958)}$$



setup unity feedback



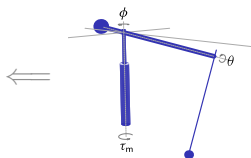
where $y = \phi$ and $u = \tau_m$, simulate with $y_{\text{ref}} = 1$ and $d = 0$

goal stabilize the system using only shaft angle measurements

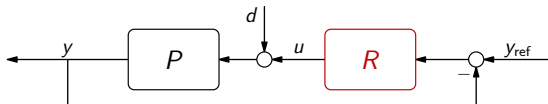
Example 2: Furuta pendulum damping

plant $P : \tau_m \mapsto \theta$ with

$$P(s) = \frac{-84.694s}{(s + 0.7387)(s^2 + 0.3864s + 37.49)}$$



setup unity feedback



where $y = \theta$ and $u = \tau_m$, simulate with $y_{\text{ref}} = 0$ and $d = 1$

goal dampen the pendulum excited by load disturbance; formally, keep

$$\left| \frac{P(j\omega)}{1 + P(j\omega)R(j\omega)} \right| \leq 1 \quad \forall \omega$$

Example 3: planar rendezvous of two satellites

Planar Clohessy–Wiltshire equations of relative motion:

$$\begin{cases} \ddot{x}(t) - 3x(t) - 2\dot{y}(t) = 0 \\ \ddot{y}(t) + 2\dot{x}(t) = u(t) \end{cases}$$

plant $P : u \mapsto \begin{bmatrix} x \\ y \end{bmatrix}$ with

$$P(s) = \begin{bmatrix} 2 \\ s^2 - 3 \end{bmatrix} \frac{1}{s^2(s^2 + 1)}$$

setup controller can measure the signals x , $\dot{x}$, y , $\dot{y}$; the system is driven by the initial conditions $x(0) = y(0) = 1$ and $\dot{x}(0) = \dot{y}(0) = 0$

goal stabilize the system $\implies$ rendezvous

Syllabus

1. review of linear algebra + some new material + matrix equations
2. state-space realizations, properties of transfer functions via realizations
3. structural properties of realizations (controllability, observability, etc)
4. state feedback, quadratic Lyapunov functions and LMIs
5. LQR optimization over finite- and infinite-horizons
6. state observers (full- and reduced-order), optimal observers
7. observer-based feedback, LQG, system design via LQG
8. handling disturbances
 - modeling disturbances: disturbance observers and internal-model control
 - worst-case disturbances: H_∞ state feedback
9. discrete-time systems and discrete design methods
10. introduction to MPC (Model Predictive Control)

Outline

Linear algebra background

Sign-definite matrices

Linear matrix equations

Outline

Linear algebra background

Sign-definite matrices

Linear matrix equations

Linear algebra, notation and terminology

A matrix $A \in \mathbb{R}^{n \times m}$ or $A \in \mathbb{C}^{n \times m}$ is an $n \times m$ table of numbers,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} =: [a_{ij}],$$

accompanied by their manipulation rules (addition, multiplication, etc). An $n \times m$ matrix A is said to be

- square if $n = m$, tall if $n > m$, and fat if $n < m$
- diagonal ($A = \text{diag}\{a_i\}$) if it is square and $a_{ij} = 0$ whenever $i \neq j$
- lower (upper) triangular if $a_{ij} = 0$ whenever $i > j$ ($i < j$)
- symmetric (or Hermitian, if over $\mathbb{C}$) if $A = A'$, where the transpose (or conjugate transpose) $A' := [\overline{a_{ji}}]$

The identity matrix $I_n := \text{diag}\{1\} \in \mathbb{R}^{n \times n}$ (if the dimension is clear, just I). It is a convention that $A^0 = I$ for every nonzero square A .

Matrix rank

The number of linearly independent rows or columns in A is called its rank and denoted $\text{rank } A$. Given $A \in \mathbb{R}^{n \times m}$,

- $\text{rank } A \leq \min\{n, m\}$; if $\text{rank } A = \min\{n, m\}$, then A is called full-rank
square full-rank matrices have $\det A \neq 0$ and referred to as nonsingular / invertible
- A is said to have full row (column) rank if $\text{rank } A = n$ ($\text{rank } A = m$)
- A has full row (column) rank iff $\eta' A = 0$ ($A\eta = 0$) $\implies \eta = 0$
- $\text{rank}(A_1 A_2) \leq \min\{\text{rank } A_1, \text{rank } A_2\}$
- if $M_1 \in \mathbb{R}^{n \times n}$, $M_2 \in \mathbb{R}^{m \times m}$ are invertible, then $\text{rank } A = \text{rank}(M_1 A M_2)$

Rank decomposition

- $\text{rank } A = r$ iff there exist full column rank $A_1 \in \mathbb{R}^{n \times r}$ and full row rank $A_2 \in \mathbb{R}^{r \times m}$ such that $A = A_1 A_2$.

for example, rank-1 matrix $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$

Matrix rank

The number of linearly independent rows or columns in A is called its rank and denoted $\text{rank } A$. Given $A \in \mathbb{R}^{n \times m}$,

- $\text{rank } A \leq \min\{n, m\}$; if $\text{rank } A = \min\{n, m\}$, then A is called full-rank
square full-rank matrices have $\det A \neq 0$ and referred to as nonsingular / invertible
- A is said to have full row (column) rank if $\text{rank } A = n$ ($\text{rank } A = m$)
- A has full row (column) rank iff $\eta' A = 0$ ($A\eta = 0$) $\implies \eta = 0$
- $\text{rank}(A_1 A_2) \leq \min\{\text{rank } A_1, \text{rank } A_2\}$
- if $M_1 \in \mathbb{R}^{n \times n}$, $M_2 \in \mathbb{R}^{m \times m}$ are invertible, then $\text{rank } A = \text{rank}(M_1 A M_2)$

Rank decomposition:

- $\text{rank } A = r$ iff there exist full column rank $A_1 \in \mathbb{R}^{n \times r}$ and full row rank $A_2 \in \mathbb{R}^{r \times m}$ such that $A = A_1 A_2$.

for example, rank-1 matrix $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \overbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}^{A_1} \overbrace{\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}}^{A_2}$

Eigenvalues & eigenvectors

Given a square $A \in \mathbb{R}^{n \times n}$, its **eigenvalues** are the solutions $\lambda \in \mathbb{C}$ to

$$\chi_A(\lambda) := \det(\lambda I - A) = \lambda^n + \chi_{n-1}\lambda^{n-1} + \cdots + \chi_1\lambda + \chi_0 = 0$$

(**characteristic equation**). There are n (not necessarily different) eigenvalues of A and $\text{spec}(A)$ denotes their set.

- if $\lambda_i \in \text{spec}(A)$ and $\text{Im } \lambda_i \neq 0$, then $\bar{\lambda}_i \in \text{spec}(A)$ too
- if $\lambda_i \in \text{spec}(A)$, then $\lambda_i t \in \text{spec}(At)$ for all $t \in \mathbb{R}$
- if $A = A'$, then all eigenvalues are real, $\text{spec}(A) \subset [\lambda_{\min}(A), \lambda_{\max}(A)]$

Right / left **eigenvectors** of $\lambda_i \in \text{spec}(A)$ are any $\eta_i, \tilde{\eta}_i \in \mathbb{C}^n \setminus \{0\}$ such that

$$(\lambda_i I - A)\eta_i = 0 \quad \text{and} \quad \tilde{\eta}_i'(\lambda_i I - A) = 0,$$

respectively (equivalently, $A\eta_i = \lambda_i\eta_i$ and $\tilde{\eta}_i'A = \lambda_i\tilde{\eta}_i'$). Note that

$$A^k\eta_i = \eta_i\lambda_i^k \quad \text{and} \quad \tilde{\eta}_i'A^k = \lambda_i^k\tilde{\eta}_i', \quad \forall k \in \mathbb{N}$$

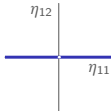
so $\lambda_i \in \text{spec}(A) \implies \lambda_i^k \in \text{spec}(A^k)$, keeping left and right eigenvectors.

Eigenvalues & eigenvectors: example

If $A = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}$, then the characteristic equation

$$\begin{aligned}\chi_A(\lambda) &= \det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}\right) = \det\begin{bmatrix} \lambda + 1 & -1 \\ 0 & \lambda - 2 \end{bmatrix} \\ &= (\lambda + 1)(\lambda - 2) = \lambda^2 - \lambda - 2 = 0\end{aligned}$$

Hence, $\text{spec}(A) = \{-1, 2\}$ (i.e. $\lambda_1 = -1$ and $\lambda_2 = 2$) and eigenvectors are $\lambda_1 = -1$:

$$\begin{bmatrix} 0 & -1 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} \eta_{11} \\ \eta_{12} \end{bmatrix} = 0 \iff \eta_{12} = 0 \iff$$


$\eta_{11} = 1$ yields $\eta_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (normalized).

$\lambda_2 = 2$:

$$\begin{bmatrix} 3 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_{21} \\ \eta_{22} \end{bmatrix} = 0 \iff \eta_{22} = 3\eta_{21} \iff$$


$\eta_{21} = 1/\sqrt{10}$ yields $\eta_2 = \begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{bmatrix}$ (normalized).

Eigenvalue multiplicities

Given $A \in \mathbb{R}^{n \times n}$ and $\lambda_i \in \text{spec}(A)$,

- algebraic multiplicity of λ_i is its multiplicity in the polynomial $\chi_A(\lambda)$
- geometric multiplicity of λ_i is $n - \text{rank}(\lambda_i I - A)$

They need not coincide. If $\exists \lambda_i \in \text{spec}(A)$ such that its algebraic multiplicity is larger than geometric one, then A is said to be *defective*.

Example 1: if $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then $\chi_A(s) = (s - 1)^2$ and

$$n - \text{rank}(\lambda_i I - A) = 2 - \text{rank}\left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right) = 2$$

i.e. both algebraic and geometric multiplicity of $\lambda_i = 1$ are 2.

Example 2: if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, then $\chi_A(s) = (s - 1)^2$ as well, but

$$n - \text{rank}(\lambda_i I - A) = 2 - \text{rank}\left(\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}\right) = 1$$

i.e. the algebraic multiplicity of $\lambda_i = 1$ is 2, whereas the geometric one is 1.
Hence, this matrix is defective.

Eigenvalue multiplicities

Given $A \in \mathbb{R}^{n \times n}$ and $\lambda_i \in \text{spec}(A)$,

- algebraic multiplicity of λ_i is its multiplicity in the polynomial $\chi_A(\lambda)$
- geometric multiplicity of λ_i is $n - \text{rank}(\lambda_i I - A)$

They need not coincide. If $\exists \lambda_i \in \text{spec}(A)$ such that its algebraic multiplicity is larger than geometric one, then A is said to be *defective*.

Example 1: if $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, then $\chi_A(s) = (s - 1)^2$ and

$$n - \text{rank}(\lambda_i I - A) = 2 - \text{rank}\left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right) = 2$$

i.e. both algebraic and geometric multiplicity of $\lambda_i = 1$ are 2.

Example 2: if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, then $\chi_A(s) = (s - 1)^2$ as well, but

$$n - \text{rank}(\lambda_i I - A) = 2 - \text{rank}\left(\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}\right) = 1$$

i.e. the algebraic multiplicity of $\lambda_i = 1$ is 2, whereas the geometric one is 1. Hence, this matrix is defective.

Similar matrices

Let $A \in \mathbb{R}^{n \times n}$ and $T \in \mathbb{R}^{n \times n}$ with $\det T \neq 0$. The matrices

$$A \quad \text{and} \quad TAT^{-1}$$

are said to be **similar** and $A \mapsto TAT^{-1}$ is called a similarity transformation. Similarity transformations do not affect characteristic equations:

$$\begin{aligned}\chi_{TAT^{-1}}(s) &= \det(\lambda I - TAT^{-1}) = \det(T(\lambda I - A)T^{-1}) \\ &= \det(T) \det(\lambda I - A) \det(T^{-1}) = \det(\lambda I - A) = \chi_A(s).\end{aligned}$$

Some definitions / facts:

- $A \in \mathbb{R}^{n \times n}$ is said to be **diagonalizable** if there is nonsingular $T \in \mathbb{R}^{n \times n}$ such that $TAT^{-1} = \text{diag}\{\lambda_i\}$, with all $\lambda_i \in \text{spec}(A)$
- A is diagonalizable iff it is not defective
- $A \in \mathbb{R}^{n \times n}$ is said to be **unitarily diagonalizable** if there exists $T \in \mathbb{R}^{n \times n}$ such that $T^{-1} = T'$ and $TAT' = \text{diag}\{\lambda_i\}$, with all $\lambda_i \in \text{spec}(A)$
- symmetric $A = A'$ is always unitarily diagonalizable

Diagonalization via eigenvectors

Let A be diagonalizable and e_i denote the standard basis of $\mathbb{R}^n$. We have:

$$TAT^{-1} = \text{diag}\{\lambda_i\} \iff AT^{-1} = T^{-1}\text{diag}\{\lambda_i\} \iff AT^{-1}e_i = \lambda_i T^{-1}e_i$$

meaning that the i th column of T^{-1} , $T^{-1}e_i = \eta_i$ (the i th eigenvector).

Example: Return to $A = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}$. Its eigenvalues are

— $\lambda_1 = -1$, with $\eta_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

— $\lambda_2 = 2$, with $\eta_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Choose

$$T^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} \implies T = \begin{bmatrix} 1 & -1/3 \\ 0 & 1/3 \end{bmatrix}.$$

In this case,

$$TAT^{-1} = \begin{bmatrix} 1 & -1/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Diagonalization via eigenvectors

Let A be diagonalizable and e_i denote the standard basis of $\mathbb{R}^n$. We have:

$$TAT^{-1} = \text{diag}\{\lambda_i\} \iff AT^{-1} = T^{-1}\text{diag}\{\lambda_i\} \iff AT^{-1}e_i = \lambda_i T^{-1}e_i$$

meaning that the i th column of T^{-1} , $T^{-1}e_i = \eta_i$ (the i th eigenvector).

Example: Return to $A = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix}$. Its eigenvalues are

- $\lambda_1 = -1$, with $\eta_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$,
- $\lambda_2 = 2$, with $\eta_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Choose

$$T^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} \implies T = \begin{bmatrix} 1 & -1/3 \\ 0 & 1/3 \end{bmatrix}.$$

In this case,

$$TAT^{-1} = \begin{bmatrix} 1 & -1/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Cayley–Hamilton

In essence: each square matrix satisfies its own characteristic equation:

$$\chi_A(A) := A^n + \chi_{n-1}A^{n-1} + \cdots + \chi_1A + \chi_0I_n = 0.$$

Important consequences:

- A^k for every $k \geq n$ is a linear combination of A^i for $i \in \mathbb{Z}_{0..n-1}$, like

$$\begin{aligned} A^n &= -\chi_{n-1}A^{n-1} - \cdots - \chi_1A - \chi_0I_n \\ A^{n+1} &= -\chi_{n-1}A^n - \cdots - \chi_1A^2 - \chi_0A \\ &= \chi_{n-1}(\chi_{n-1}A^{n-1} + \cdots + \chi_1A + \chi_0I_n) - \cdots - \chi_1A^2 - \chi_0A \\ &= (\chi_{n-1}^2 - \chi_{n-2})A^{n-1} + \cdots + (\chi_{n-1}\chi_1 - \chi_0)A + \chi_{n-1}\chi_0I_n \\ &\vdots \end{aligned}$$

- A^{-1} , if exists, is also a linear combination of A^i , $i \in \mathbb{Z}_{0..n-1}$:

$$A^{-1} = -\frac{1}{\chi_0}(\chi_1I + \cdots + \chi_{n-1}A^{n-2} + A^{n-1}).$$

Cayley–Hamilton: example

Consider again

$$A = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix} \implies \chi_A(\lambda) = \lambda^2 - \lambda - 2$$

(i.e. $\chi_1 = -1$ and $\chi_0 = -2$). Hence,

$$\chi_A(A) = A^2 - A - 2I_2 = \begin{bmatrix} 1 & 1 \\ 0 & 4 \end{bmatrix} - \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

and

$$A^2 = A + 2I_2 = \begin{bmatrix} -1 & 1 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 4 \end{bmatrix}$$

$$A^3 = A^2 + 2A = A + 2I_2 + 2A = 3A + 2I_2 = \begin{bmatrix} -1 & 3 \\ 0 & 8 \end{bmatrix}$$

$$A^4 = 3A^2 + 2A = 3A + 6I_2 + 2A = 5A + 6I_2 = \begin{bmatrix} 1 & 5 \\ 0 & 16 \end{bmatrix}$$

$\vdots$

Matrix functions

Let $f(x) = \sum_{i=0}^{\infty} f_i x^i$ be analytic. Given a square A , its function

$$f(A) := \sum_{i=0}^{\infty} f_i A^i.$$

Properties:

- $f(A)\eta_i = f(\lambda_i)\eta_i$ for all eigenvalue-eigenvector pairs (λ_i, η_i)
- $f(TAT^{-1}) = Tf(A)T^{-1}$ for any analytic function $f(x)$
because $(TAT^{-1})^i = (TAT^{-1})(TAT^{-1})\cdots(TAT^{-1}) = TAA\cdots AT^{-1} = TA^i T^{-1}$
- $f(\text{diag}\{\lambda_i\}) = \text{diag}\{f(\lambda_i)\}$
consequently, $f(A) = T \text{diag}\{f(\lambda_i)\} T^{-1}$ whenever A is diagonalizable

By Cayley–Hamilton, we know that $\exists g, f \in \mathbb{Z}_{0, n-1}$ such that

$$f(A) = \sum_{i=0}^{n-1} g_i A^i,$$

with $f(\lambda_j) = g(\lambda_j)$ for all $\lambda_j \in \text{spec}(A)$, where $g(x) = \sum_{i=0}^{n-1} g_i x^i$.

Matrix functions

Let $f(x) = \sum_{i=0}^{\infty} f_i x^i$ be analytic. Given a square A , its function

$$f(A) := \sum_{i=0}^{\infty} f_i A^i.$$

Properties:

- $f(A)\eta_i = f(\lambda_i)\eta_i$ for all eigenvalue-eigenvector pairs (λ_i, η_i)
- $f(TAT^{-1}) = Tf(A)T^{-1}$ for any analytic function $f(x)$
because $(TAT^{-1})^i = (TAT^{-1})(TAT^{-1})\cdots(TAT^{-1}) = TAA\cdots AT^{-1} = TA^i T^{-1}$
- $f(\text{diag}\{\lambda_i\}) = \text{diag}\{f(\lambda_i)\}$
consequently, $f(A) = T \text{diag}\{f(\lambda_i)\} T^{-1}$ whenever A is diagonalizable

By Cayley–Hamilton, we know that $\exists g_i, i \in \mathbb{Z}_{0..n-1}$ such that

$$f(A) = \sum_{i=0}^{n-1} g_i A^i,$$

with $f(\lambda_j) = g(\lambda_j)$ for all $\lambda_j \in \text{spec}(A)$, where $g(x) = \sum_{i=0}^{n-1} g_i x^i$.

Matrix functions: exponential

The matrix exponential is defined (here $t \in \mathbb{R}$, we shall need it later on) as

$$\exp(At) = e^{At} := I + At + \frac{1}{2!}(At)^2 + \frac{1}{3!}(At)^3 + \dots$$

Properties:

- $(e^{At})' = e^{A't}$
- if $A = 0$ or $t = 0$, then $e^{At} = I$
- $A = \text{diag}\{a_i\} \implies e^{At} = \text{diag}\{e^{a_i t}\}$
- if A is lower / upper triangular, then so is e^{At}
- $e^{At}e^{Bt} = e^{(A+B)t} \iff AB = BA$
 - hence: $e^{\lambda t}e^{At} = e^{(\lambda I + A)t}$, $e^{At}e^{A\tau} = e^{A(t+\tau)}$, & $e^{At}e^{-At} = I = e^{At}(e^{At})^{-1}$
- $\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A$
- $\lambda \in \text{spec}(A) \implies e^{\lambda t} \in \text{spec}(e^{At})$
 - if $\lambda = \pm j \in \text{spec}(A)$, then $e^{\pm jt} \in \text{spec}(e^{At})$, but $e^{jt} = e^{-jt}$, $\forall t = \pi + 2\pi k$, $k \in \mathbb{Z}$
- Laplace transform: $\mathfrak{L}\{e^{At}\mathbb{1}(t)\} = (sI - A)^{-1}$ aka resolvent

Block matrices

Let $A \in \mathbb{R}^{n \times m}$ and let $\sum_{i=1}^{\nu} n_i = n$ and $\sum_{i=1}^{\mu} m_i = m$ for some $n_i, m_i \in \mathbb{N}$. Then A can be formally split as

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1\mu} \\ A_{21} & A_{22} & \cdots & A_{2\mu} \\ \vdots & \vdots & \ddots & \vdots \\ A_{\nu 1} & A_{\nu 2} & \cdots & A_{\nu \mu} \end{bmatrix} = [A_{ij}]$$

for $A_{ij} = [a_{ij,kl}] \in \mathbb{R}^{n_i \times m_j}$ with entries $a_{ij,kl} = a_{pq}$, where $p = \sum_{r=1}^{i-1} n_r + k$ and $q = \sum_{r=1}^{j-1} m_r + l$. For example, we may present

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

for $A_{11} = \begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$, $A_{13} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$, $A_{21} = \begin{bmatrix} 9 & 10 \end{bmatrix}$, $A_{22} = 11$, $A_{23} = 12$. This split is not unique and is normally done merely for convenience.

Block matrices (contd)

We say that a block matrix $A = [A_{ij}]$ is

- block-diagonal ($A = \text{diag}\{A_i\}$) if $\nu = \mu$ and $A_{ij} = 0$ whenever $i \neq j$
- block lower (upper) triangular if $A_{ij} = 0$ whenever $i > j$ ($i < j$)

Operations on block matrices are performed in terms of their blocks, like

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}' = \begin{bmatrix} A_{11}' & A_{21}' \\ A_{12}' & A_{22}' \end{bmatrix},$$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix},$$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

Sum, product, and (if all diagonal blocks are square) inverse and functions of block-diagonal/triangular matrices remain block-diagonal/triangular.

Eigenvalues & eigenvectors of block-triangular matrices

Let $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ with square A_{11} and A_{22} . Because

$$\begin{bmatrix} \lambda I - A_{11} & -A_{12} \\ 0 & \lambda I - A_{22} \end{bmatrix} \begin{bmatrix} \eta \\ 0 \end{bmatrix} = \begin{bmatrix} (\lambda I - A_{11})\eta \\ 0 \end{bmatrix},$$

every eigenvalue of A_{11} is that of A (its right eigenvector is $\begin{bmatrix} \eta_i \\ 0 \end{bmatrix}$). Because

$$\begin{bmatrix} 0 & \tilde{\eta}' \end{bmatrix} \begin{bmatrix} \lambda I - A_{11} & -A_{12} \\ 0 & \lambda I - A_{22} \end{bmatrix} = \begin{bmatrix} 0 & \tilde{\eta}'(\lambda I - A_{22}) \end{bmatrix},$$

every eigenvalue of A_{22} is that of A (its left eigenvector is $\begin{bmatrix} 0 & \tilde{\eta}'_i \end{bmatrix}$). Thus

- λ_i is an eigenvalue of A iff it is an eigenvalue of A_{11} or A_{22} .

The same arguments

- work for block-lower-triangular and block-diagonal matrices too.

Block LDU and UDL decompositions

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}.$$

If $\det A_{11} \neq 0$, then

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \overbrace{\begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix}}^{\text{nonsingular}} \overbrace{\begin{bmatrix} A_{11} & 0 \\ 0 & \Delta_{11} \end{bmatrix}}^{\text{block-diagonal}} \overbrace{\begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix}}^{\text{nonsingular}},$$

where $\Delta_{11} := A_{22} - A_{21}A_{11}^{-1}A_{12}$ is the **Schur complement of A_{11}** in A . Note that if A_{22} is also square, then $\det A = \det A_{11} \det \Delta_{11}$ then.

Likewise, if $\det A_{22} \neq 0$, then

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Delta_{22} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{bmatrix},$$

where $\Delta_{22} := A_{11} - A_{12}A_{22}^{-1}A_{21}$ is the **Schur complement of A_{22}** in A .

Block LDU decomposition: example

Let

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \overbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}}^{\text{partition 1}} = \overbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}}^{\text{partition 2}}$$

Partition 1: $A_{11} = 1 \implies \Delta_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} 1^{-1} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and then

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

Partition 2: $A_{11} = I_2 \implies \Delta_{11} = 2 - \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0$ and then

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

Matrix calculus

The derivative of $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ by its vector argument $x \in \mathbb{R}^n$,

$$\frac{\partial f(x)}{\partial x} = \left[\frac{\partial f_i(x)}{\partial x_j} \right] \in \mathbb{R}^{m \times n} \quad \text{for every } x$$

i.e. it is a matrix-valued function, $\mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$, of x .

Example: Let $f(x) = x'Px \in \mathbb{R}$ for some symmetric $P \in \mathbb{R}^{n \times n}$. Its derivative $\partial f(x)/\partial x \in \mathbb{R}^{1 \times n}$ is a row vector, whose k th element is

$$\begin{aligned} \frac{\partial f(x)}{\partial x_k} &= \frac{\partial}{\partial x_k} \sum_{i=1}^n \sum_{j=1}^n p_{ij} x_i x_j = \frac{\partial p_{kk} x_k^2}{\partial x_k} + \frac{\partial}{\partial x_k} \sum_{\substack{i=1 \\ i \neq k}}^n p_{ik} x_i x_k + \frac{\partial}{\partial x_k} \sum_{\substack{j=1 \\ j \neq k}}^n p_{kj} x_k x_j \\ &= 2p_{kk} x_k + \sum_{i=1, i \neq k}^n p_{ik} x_i + \sum_{j=1, j \neq k}^n p_{kj} x_j = 2 \sum_{i=1}^n x_i p_{ik} \end{aligned}$$

with the use of $p_{ij} = p_{ji}$ (because $P = P'$). Hence, $\frac{\partial}{\partial x} x'Px = 2x'P$.

Matrix calculus

The derivative of $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ by its vector argument $x \in \mathbb{R}^n$,

$$\frac{\partial f(x)}{\partial x} = \left[\frac{\partial f_i(x)}{\partial x_j} \right] \in \mathbb{R}^{m \times n} \quad \text{for every } x$$

i.e. it is a matrix-valued function, $\mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$, of x .

Example: Let $f(x) = x'Px \in \mathbb{R}$ for some symmetric $P \in \mathbb{R}^{n \times n}$. Its derivative $\partial f(x)/\partial x \in \mathbb{R}^{1 \times n}$ is a row vector, whose k th element is

$$\begin{aligned} \frac{\partial f(x)}{\partial x_k} &= \frac{\partial}{\partial x_k} \sum_{i=1}^n \sum_{j=1}^n p_{ij} x_i x_j = \frac{\partial p_{kk} x_k^2}{\partial x_k} + \frac{\partial}{\partial x_k} \sum_{\substack{i=1 \\ i \neq k}}^n p_{ik} x_i x_k + \frac{\partial}{\partial x_k} \sum_{\substack{j=1 \\ j \neq k}}^n p_{kj} x_k x_j \\ &= 2p_{kk} x_k + \sum_{i=1, i \neq k}^n p_{ik} x_i + \sum_{j=1, j \neq k}^n p_{kj} x_j = 2 \sum_{i=1}^n x_i p_{ik} \end{aligned}$$

with the use of $p_{ij} = p_{ji}$ (because $P = P'$). Hence, $\frac{\partial}{\partial x} x' P x = 2x' P$.

Outline

Linear algebra background

Sign-definite matrices

Linear matrix equations

Sign-definiteness of matrices

A symmetric matrix $A = A' \in \mathbb{R}^{n \times n}$ is said to be

- positive definite ($A > 0$) if $x'Ax > 0$ for all $0 \neq x \in \mathbb{R}^n$
- positive semi-definite ($A \geq 0$) if $x'Ax \geq 0$ for all $x \in \mathbb{R}^n$
obviously, $A > 0 \implies A \geq 0$ as well, so saying that $A \geq 0$ does not rule out $A > 0$
- negative definite ($A < 0$) if $x'Ax < 0$ for all $0 \neq x \in \mathbb{R}^n$ (or if $-A > 0$)
- negative semi-definite ($A \leq 0$) if $x'Ax \leq 0$ for all $x \in \mathbb{R}^n$ (or $-A \geq 0$)

Given two matrices $A = A' \in \mathbb{R}^{n \times n}$ and $B = B' \in \mathbb{R}^{n \times n}$, we say

- $A > B$ if $A - B > 0$
- $A \geq B$ if $A - B \geq 0$
- $A < B$ if $B - A > 0$ (equivalently, $A - B < 0$)
- $A \leq B$ if $B - A \geq 0$ (equivalently, $A - B \leq 0$)

Sign-definiteness of matrices

A symmetric matrix $A = A' \in \mathbb{R}^{n \times n}$ is said to be

- positive definite ($A > 0$) if $x'Ax > 0$ for all $0 \neq x \in \mathbb{R}^n$
- positive semi-definite ($A \geq 0$) if $x'Ax \geq 0$ for all $x \in \mathbb{R}^n$
obviously, $A > 0 \implies A \geq 0$ as well, so saying that $A \geq 0$ does not rule out $A > 0$
- negative definite ($A < 0$) if $x'Ax < 0$ for all $0 \neq x \in \mathbb{R}^n$ (or if $-A > 0$)
- negative semi-definite ($A \leq 0$) if $x'Ax \leq 0$ for all $x \in \mathbb{R}^n$ (or $-A \geq 0$)

Given two matrices $A = A' \in \mathbb{R}^{n \times n}$ and $B = B' \in \mathbb{R}^{n \times n}$, we say

- $A > B$ if $A - B > 0$
- $A \geq B$ if $A - B \geq 0$
- $A < B$ if $B - A > 0$ (equivalently, $A - B < 0$)
- $A \leq B$ if $B - A \geq 0$ (equivalently, $A - B \leq 0$)

Sign-definite matrices: example

Let

$$A = \begin{bmatrix} \alpha & -1 \\ -1 & 1 \end{bmatrix}$$

for some $\alpha \in \mathbb{R}$. The quadratic form

$$\begin{aligned} x'Ax &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \alpha & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \alpha x_1^2 - 2x_1x_2 + x_2^2 \\ &= (\alpha - 1)x_1^2 + (x_1 - x_2)^2 \end{aligned}$$

Three cases:

$\alpha > 1 \implies x'Ax > 0$ unless $x_1 = x_2 = 0$, for which $x'Ax = 0 \implies A > 0$

$\alpha = 1 \implies x'Ax > 0$ unless $x_1 = x_2$, for which $x'Ax = 0 \implies A \geq 0$

$\alpha < 1 \implies x'Ax$ might be both positive (e.g. if $x_1 = 0 \neq x_2$)
and negative (e.g. if $x_1 = x_2 \neq 0) \implies A \not\geq 0$

Sign-definite matrices: example

Let

$$A = \begin{bmatrix} \alpha & -1 \\ -1 & 1 \end{bmatrix}$$

for some $\alpha \in \mathbb{R}$. The quadratic form

$$\begin{aligned} x'Ax &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \alpha & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \alpha x_1^2 - 2x_1x_2 + x_2^2 \\ &= (\alpha - 1)x_1^2 + (x_1 - x_2)^2 \end{aligned}$$

Three cases:

$$\alpha > 1 \implies x'Ax > 0 \text{ unless } x_1 = x_2 = 0, \text{ for which } x'Ax = 0 \implies A > 0$$

$$\alpha = 1 \implies x'Ax > 0 \text{ unless } x_1 = x_2, \text{ for which } x'Ax = 0 \implies A \geq 0$$

$$\begin{aligned} \alpha < 1 \implies x'Ax \text{ might be both positive (e.g. if } x_1 = 0 \neq x_2) \\ \text{and negative (e.g. if } x_1 = x_2 \neq 0) \implies A \not\geq 0 \end{aligned}$$

Properties of sign-definite matrices

Let $A = A' \in \mathbb{R}^{n \times n}$.

1. $A > 0 \implies \det A \neq 0$

$\det A = 0 \implies \exists x \neq 0$ such that $Ax = 0$ and then $x'Ax = 0 \implies A \not> 0$

2. if $\det N \neq 0$, then $A > 0 \iff N'AN > 0$ and $A \geq 0 \iff N'AN \geq 0$
 $x'Ax = y'Ny$ for $y = Nx$, with $y = 0 \iff x = 0$

3. $A > 0$ (≥ 0) iff all its eigenvalues are positive (nonnegative)
 by Property 2 and unitary diagonalization of A

4. if $A \geq 0$, then $\lambda_{\min}(A)\|x\|^2 \leq x'Ax \leq \lambda_{\max}(A)\|x\|^2$ for all $x \in \mathbb{R}^n$
 by unitary diagonalization of A , both bounds are tight (i.e. equal for some x)

5. $A > 0 \iff A^{-1} > 0$

$(A^{-1})' = A^{-1} \implies A^{-1} = (A')^{-1} = (A^{-1})'AA^{-1}$, then by Property 3 with $N = A^{-1}$

6. if $A > 0$ (≥ 0) and $B \geq 0$, then $A + B > 0$ (≥ 0) as well
 $(x'Ax + Bx) = x'(A+B)x \geq x'Ax > 0$

7. $A > B > C \implies A > C$ and $A \geq B \geq C \implies A \geq C$
 $(x'Ax - Cx) = x'(A - C)x \geq x'(B - C)x \geq 0$

Properties of sign-definite matrices

Let $A = A' \in \mathbb{R}^{n \times n}$.

$$1. A > 0 \implies \det A \neq 0$$

$$\det A = 0 \implies \exists x \neq 0 \text{ such that } Ax = 0 \text{ and then } x'Ax = 0 \implies A \not> 0$$

$$2. \text{ if } \det N \neq 0, \text{ then } A > 0 \iff N'AN > 0 \text{ and } A \geq 0 \iff N'AN \geq 0$$

$$x'N'ANx = y'Ay \text{ for } y := Nx, \text{ with } y = 0 \iff x = 0$$

$$3. A > 0 (\geq 0) \text{ iff all its eigenvalues are positive (nonnegative)}$$

$$4. \text{ if } A \geq 0, \text{ then } \lambda_{\min}(A)\|x\|^2 \leq x'Ax \leq \lambda_{\max}(A)\|x\|^2 \text{ for all } x \in \mathbb{R}^n$$

$$\text{necessary eigenvalue condition of } A, \text{ both bounds are sharp (i.e. } \exists x \text{ such that the bounds are attained)}$$

$$5. A > 0 \iff A^{-1} > 0$$

$$(A^{-1})^{-1} = A \implies A^{-1} = (A^{-1})'AA^{-1}, \text{ verify Property 3 with } B = A^{-1}$$

$$6. \text{ if } A > 0 (\geq 0) \text{ and } B \geq 0, \text{ then } A+B > 0 (\geq 0) \text{ as well}$$

$$A+B > 0 \implies A > 0$$

$$7. A > B > C \implies A > C \text{ and } A \geq B \geq C \implies A \geq C$$

$$A > 0, B > 0 \implies A^{-1} > 0, B^{-1} > 0$$

Properties of sign-definite matrices

Let $A = A' \in \mathbb{R}^{n \times n}$.

- $A > 0 \implies \det A \neq 0$

$$\det A = 0 \implies \exists x \neq 0 \text{ such that } Ax = 0 \text{ and then } x'Ax = 0 \implies A \not> 0$$

- if $\det N \neq 0$, then $A > 0 \iff N'AN > 0$ and $A \geq 0 \iff N'AN \geq 0$

$$x'N'ANx = y'Ay \text{ for } y := Nx, \text{ with } y = 0 \iff x = 0$$

- $A > 0$ (≥ 0) iff all its eigenvalues are positive (nonnegative)

by Property 2 and unitary diagonalization of A

- if $A \geq 0$, then $\lambda_{\min}(A)\|x\|^2 \leq x'Ax \leq \lambda_{\max}(A)\|x\|^2$ for all $x \in \mathbb{R}^n$

$$\lambda_{\min}(A) = \min_{\|x\|=1} x'Ax \text{ and } \lambda_{\max}(A) = \max_{\|x\|=1} x'Ax \text{ (Rayleigh's theorem)}$$

- $A > 0 \iff A^{-1} > 0$

$$(A^{-1})' = (A')^{-1} = A^{-1} \text{ (use Property 2 and } A^{-1}A = I)$$

- if $A > 0$ (≥ 0) and $B \geq 0$, then $A + B > 0$ (≥ 0) as well

$$(A + B)' = A' + B' = A + B \text{ (use Property 2 and } A, B \text{ are symmetric)}$$

- $A > B > C \implies A > C$ and $A \geq B \geq C \implies A \geq C$

$$A > B \text{ and } B > C \implies A > C \text{ (use Property 2 and } A - B > 0, B - C > 0)$$

Properties of sign-definite matrices

Let $A = A' \in \mathbb{R}^{n \times n}$.

- $A > 0 \implies \det A \neq 0$

$$\det A = 0 \implies \exists x \neq 0 \text{ such that } Ax = 0 \text{ and then } x'Ax = 0 \implies A \not> 0$$

- if $\det N \neq 0$, then $A > 0 \iff N'AN > 0$ and $A \geq 0 \iff N'AN \geq 0$

$$x'N'ANx = y'Ay \text{ for } y := Nx, \text{ with } y = 0 \iff x = 0$$

- $A > 0$ (≥ 0) iff all its eigenvalues are positive (nonnegative)

by Property 2 and unitary diagonalization of A

- if $A \geq 0$, then $\lambda_{\min}(A)\|x\|^2 \leq x'Ax \leq \lambda_{\max}(A)\|x\|^2$ for all $x \in \mathbb{R}^n$

via unitary diagonalization of A ; both bounds are tight (i.e. equalities are attained)

- $A > 0 \iff A^{-1} > 0$

$$A > 0 \implies \lambda_{\min}(A) \geq \lambda_{\min}(A^{-1})^{-1} \implies \lambda_{\min}(A^{-1}) \geq \lambda_{\min}(A)^{-1}$$

- if $A > 0$ (≥ 0) and $B \geq 0$, then $A+B > 0$ (≥ 0) as well

- $A > B > C \implies A > C$ and $A \geq B \geq C \implies A \geq C$

Properties of sign-definite matrices

Let $A = A' \in \mathbb{R}^{n \times n}$.

1. $A > 0 \implies \det A \neq 0$

$\det A = 0 \implies \exists x \neq 0$ such that $Ax = 0$ and then $x'Ax = 0 \implies A \not> 0$

2. if $\det N \neq 0$, then $A > 0 \iff N'AN > 0$ and $A \geq 0 \iff N'AN \geq 0$

$x'N'ANx = y'Ay$ for $y := Nx$, with $y = 0 \iff x = 0$

3. $A > 0$ (≥ 0) iff all its eigenvalues are positive (nonnegative)

by Property 2 and unitary diagonalization of A

4. if $A \geq 0$, then $\lambda_{\min}(A)\|x\|^2 \leq x'Ax \leq \lambda_{\max}(A)\|x\|^2$ for all $x \in \mathbb{R}^n$

via unitary diagonalization of A ; both bounds are tight (i.e. equalities are attained)

5. $A > 0 \iff A^{-1} > 0$

$(A^{-1})' = A^{-1} \implies A^{-1} = (A^{-1})'AA^{-1}$, then by Property 2 with $N = A^{-1}$

6. if $A > 0$ (≥ 0) and $B \geq 0$, then $A + B > 0$ (≥ 0) as well

7. $A > B > C \implies A > C$ and $A \geq B \geq C \implies A \geq C$

Properties of sign-definite matrices

Let $A = A' \in \mathbb{R}^{n \times n}$.

- $A > 0 \implies \det A \neq 0$

$$\det A = 0 \implies \exists x \neq 0 \text{ such that } Ax = 0 \text{ and then } x'Ax = 0 \implies A \not> 0$$

- if $\det N \neq 0$, then $A > 0 \iff N'AN > 0$ and $A \geq 0 \iff N'AN \geq 0$

$$x'N'ANx = y'Ay \text{ for } y := Nx, \text{ with } y = 0 \iff x = 0$$

- $A > 0$ (≥ 0) iff all its eigenvalues are positive (nonnegative)

by Property 2 and unitary diagonalization of A

- if $A \geq 0$, then $\lambda_{\min}(A)\|x\|^2 \leq x'Ax \leq \lambda_{\max}(A)\|x\|^2$ for all $x \in \mathbb{R}^n$

via unitary diagonalization of A ; both bounds are tight (i.e. equalities are attained)

- $A > 0 \iff A^{-1} > 0$

$$(A^{-1})' = A^{-1} \implies A^{-1} = (A^{-1})'AA^{-1}, \text{ then by Property 2 with } N = A^{-1}$$

- if $A > 0$ (≥ 0) and $B \geq 0$, then $A + B > 0$ (≥ 0) as well

$$\text{by } x'(A + B)x = x'Ax + x'Bx$$

- $A > B > C \implies A > C$ and $A \geq B \geq C \implies A \geq C$

Properties of sign-definite matrices

Let $A = A' \in \mathbb{R}^{n \times n}$.

- $A > 0 \implies \det A \neq 0$

$$\det A = 0 \implies \exists x \neq 0 \text{ such that } Ax = 0 \text{ and then } x'Ax = 0 \implies A \not> 0$$

- if $\det N \neq 0$, then $A > 0 \iff N'AN > 0$ and $A \geq 0 \iff N'AN \geq 0$

$$x'N'ANx = y'Ay \text{ for } y := Nx, \text{ with } y = 0 \iff x = 0$$

- $A > 0$ (≥ 0) iff all its eigenvalues are positive (nonnegative)

by Property 2 and unitary diagonalization of A

- if $A \geq 0$, then $\lambda_{\min}(A)\|x\|^2 \leq x'Ax \leq \lambda_{\max}(A)\|x\|^2$ for all $x \in \mathbb{R}^n$

via unitary diagonalization of A ; both bounds are tight (i.e. equalities are attained)

- $A > 0 \iff A^{-1} > 0$

$$(A^{-1})' = A^{-1} \implies A^{-1} = (A^{-1})'AA^{-1}, \text{ then by Property 2 with } N = A^{-1}$$

- if $A > 0$ (≥ 0) and $B \geq 0$, then $A + B > 0$ (≥ 0) as well

$$\text{by } x'(A + B)x = x'Ax + x'Bx$$

- $A > B > C \implies A > C$ and $A \geq B \geq C \implies A \geq C$

$$\text{by } x'(A - C)x = x'(A - B)x + x'(B - C)x$$

Properties of sign-definite matrices (contd)

8. $A > 0 \iff \exists M \in \mathbb{R}^{n \times n}$ such that $\det M \neq 0$ and $A = M'M$

Cholesky decomposition if M is upper triangular with positive diagonal elements

9. $(A \geq 0) \wedge (\text{rank } A = r) \iff \exists M \in \mathbb{R}^{r \times n}$ such that

$$\text{rank } M = r \quad \text{and} \quad A = M'M;$$

10. $A \geq 0 \implies \exists M = M' \in \mathbb{R}^{n \times n}$ such that

$$M \geq 0, \quad \text{rank } M = \text{rank } A \quad \text{and} \quad A = M^2 = MM,$$

such M is unique and dubbed the matrix square root of A , $M = A^{1/2}$

11. if $A \geq 0$, then $x'Ax = 0 \implies Ax = 0$

$$\text{by Property 9, } x'Ax = 0 \iff \|Mx\|^2 = 0 \iff Mx = 0 \iff M'Mx = 0$$

12. $A > I \ (A \geq I) \iff A^{-1} < I \ (A^{-1} \leq I)$

$$\text{by Properties 2 and 10, } A > I \iff A^{-1/2}AA^{-1/2} > A^{-1/2}IA^{-1/2} \iff I > A^{-1} \dots$$

13. if $A > 0$, $B > 0$, then $A > B \ (A \geq B) \iff A^{-1} < B^{-1} \ (A^{-1} \leq B^{-1})$

$$\text{by Properties 10, 2, 12, } A \geq B \iff B^{-1/2}AB^{-1/2} \geq I \iff B^{1/2}A^{-1}B^{1/2} \leq I \dots$$

Sign-definite block matrices

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A'_{12} & A_{22} \end{bmatrix} \text{ for } A_{11} = A'_{11} \text{ and } A_{22} = A'_{22}.$$

Lemma 1.1

$A > 0$ only if both $A_{11} > 0$ and $A_{22} > 0$.

Proof: Assume the contrary, that $A > 0$, yet $A_{11} \not> 0$. Hence, $\exists x_1 \neq 0$ such that $x'_1 A_{11} x_1 \not> 0$. Hence,

$$\begin{bmatrix} x'_1 & 0 \end{bmatrix} A \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = x'_1 A_{11} x_1 \not> 0 \implies A \not> 0,$$

which is a contradiction. Similar arguments work for A_{22} . □

Lemma 1.2

$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} > 0$ iff both $A_{11} > 0$ and $A_{22} > 0$.

Proof: Follows by $\begin{bmatrix} x'_1 & x'_2 \end{bmatrix} A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x'_1 A_{11} x_1 + x'_2 A_{22} x_2$. □

Sign-definite block matrices

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A'_{12} & A_{22} \end{bmatrix} \text{ for } A_{11} = A'_{11} \text{ and } A_{22} = A'_{22}.$$

Lemma 1.1

$A > 0$ only if both $A_{11} > 0$ and $A_{22} > 0$.

Proof: Assume the contrary, that $A > 0$, yet $A_{11} \not> 0$. Hence, $\exists x_1 \neq 0$ such that $x'_1 A_{11} x_1 \not> 0$. Hence,

$$\begin{bmatrix} x'_1 & 0 \end{bmatrix} A \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = x'_1 A_{11} x_1 \not> 0 \implies A \not> 0,$$

which is a contradiction. Similar arguments work for A_{22} . □

Lemma 1.2

$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} > 0$ (≥ 0) iff both $A_{11} > 0$ (≥ 0) and $A_{22} > 0$ (≥ 0).

Proof: Follows by $\begin{bmatrix} x'_1 & x'_2 \end{bmatrix} A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x'_1 A_{11} x_1 + x'_2 A_{22} x_2$. □

Sign-definite block matrices (contd)

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A'_{12} & A_{22} \end{bmatrix} \text{ for } A_{11} = A'_{11} \text{ and } A_{22} = A'_{22}.$$

Theorem 1.3

The following statements are equivalent

1. $A > 0$
2. $A_{11} > 0$ and $\Delta_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12} > 0$
3. $A_{22} > 0$ and $\Delta_{22} = A_{11} - A_{12}A_{22}^{-1}A_{21} > 0$

Proof: 1 $\iff$ 2 follows by Lemmas 1.1 and 1.2, the LDU decomposition

$$A = \begin{bmatrix} I & 0 \\ A'_{12}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & \Delta_{11} \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix},$$

and the property that $A > 0 \iff N'AN > 0$ for any nonsingular matrix N .

1 $\iff$ 3 follows by similar arguments via the UDL decomposition. $\square$

Sign-definite block matrices: example

Return to

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \overbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}}^{\text{partition 1}} = \overbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}}^{\text{partition 2}}$$

Partition 1: $A_{11} = 1 > 0$ and $\Delta_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \stackrel{?}{>} 0$

Partition 2: $A_{11} = I_2 > 0$ and $\Delta_{11} = 0 \not> 0$

way clearer!

Moral:

“right” partition could be a game changer.

Sign-definite block matrices: example

Return to

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \overbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}}^{\text{partition 1}} = \overbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}}^{\text{partition 2}}$$

Partition 1: $A_{11} = 1 > 0$ and $\Delta_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \stackrel{?}{>} 0$

Partition 2: $A_{11} = I_2 > 0$ and $\Delta_{11} = 0 \not> 0$

way clearer!

Moral:

- “right” partition could be a game changer.

Sign-definite block matrices (contd)

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A'_{12} & A_{22} \end{bmatrix} \text{ for } A_{11} = A'_{11} \text{ and } A_{22} = A'_{22}.$$

Theorem 1.4

1. if $A_{11} > 0$, then $A \geq 0 \iff \Delta_{11} = A_{22} - A_{21}A_{11}^{-1}A_{12} \geq 0$
2. if $A_{22} > 0$, then $A \geq 0 \iff \Delta_{22} = A_{11} - A_{12}A_{22}^{-1}A_{21} \geq 0$

Proof: 1 $\iff$ 3 follows by Lemma 1.2, the UDL decomposition

$$A = \begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Delta_{22} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ A_{22}^{-1}A'_{12} & I \end{bmatrix}$$

and the property that $A \geq 0 \iff N'AN \geq 0$ for any nonsingular matrix N .

1 $\iff$ 2 follows by similar arguments via the LDU decomposition. $\square$

Sign-definite block matrices: example (contd)

Return to

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

By Theorem 1.4, $A \geq 0$. Hence, we may perform the following

Rank decomposition

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} = \overbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}}^{M'} \overbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}}^M$$

Square root (could be verified via $A^{1/2}A^{1/2} = A$)

$$A^{1/2} = \frac{1}{6} \begin{bmatrix} 3 + \sqrt{3} & -3 + \sqrt{3} & 2\sqrt{3} \\ -3 + \sqrt{3} & 3 + \sqrt{3} & 2\sqrt{3} \\ 2\sqrt{3} & 2\sqrt{3} & 4\sqrt{3} \end{bmatrix} \geq 0$$

Outline

Linear algebra background

Sign-definite matrices

Linear matrix equations

Matrix Sylvester equation / Matlab: `lyap(A1,A2,Q)`

Given $A_1 \in \mathbb{R}^{p \times p}$, $A_2 \in \mathbb{R}^{m \times m}$, and $Q \in \mathbb{R}^{p \times m}$, find $X \in \mathbb{R}^{p \times m}$ such that

$$A_1 X + X A_2 + Q = 0.$$

X exists for all Q and unique iff $\lambda + \mu \neq 0$, $\forall \lambda \in \text{spec}(A_1), \mu \in \text{spec}(A_2)$.

Example: Let $A_1 = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} \alpha & 0 \\ 1 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$. In this case

$$\begin{aligned} - \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} &= \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} + \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (\alpha - 1)x_{11} + x_{21} + x_{12} & x_{22} - x_{12} \\ (\alpha + 1)x_{21} + x_{22} & x_{22} \end{bmatrix} \end{aligned}$$

so $x_{22} = -q_{22}$, $x_{12} = q_{12} - q_{22}$,

$$(\alpha + 1)x_{21} = q_{22} - q_{21}, \quad \text{and} \quad (\alpha - 1)x_{11} = q_{22} - q_{12} - x_{21} - q_{11}.$$

Solvable $\forall Q$ iff $\alpha \notin \{-1, 1\}$. Otherwise, depends on Q and is not unique.

Matrix Sylvester equation / Matlab: `lyap(A1,A2,Q)`

Given $A_1 \in \mathbb{R}^{p \times p}$, $A_2 \in \mathbb{R}^{m \times m}$, and $Q \in \mathbb{R}^{p \times m}$, find $X \in \mathbb{R}^{p \times m}$ such that

$$A_1 X + X A_2 + Q = 0.$$

X exists for all Q and unique iff $\lambda + \mu \neq 0$, $\forall \lambda \in \text{spec}(A_1), \mu \in \text{spec}(A_2)$.

Example: Let $A_1 = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} \alpha & 0 \\ 1 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$. In this case

$$\begin{aligned} - \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} &= \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} + \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} \alpha & 0 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (\alpha - 1)x_{11} + x_{21} + x_{12} & x_{22} - x_{12} \\ (\alpha + 1)x_{21} + x_{22} & x_{22} \end{bmatrix} \end{aligned}$$

so $x_{22} = -q_{22}$, $x_{12} = q_{12} - q_{22}$,

$$(\alpha + 1)x_{21} = q_{22} - q_{21}, \quad \text{and} \quad (\alpha - 1)x_{11} = q_{22} - q_{12} - x_{21} - q_{11}.$$

Solvable $\forall Q$ iff $\alpha \notin \{-1, 1\}$. Otherwise, depends on Q and is not unique.

Sylvester equation and Roth's removal rule

Matrices

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

are similar iff the Sylvester equation $A_{11}X - XA_{22} + A_{12} = 0$ is solvable.

Proof (if part): If $A_{11}X - XA_{22} + A_{12} = 0$ is solvable, then

$$\begin{aligned} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}^{-1} &= \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{11}X - XA_{22} + A_{12} \\ 0 & A_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \end{aligned}$$

This is what we call *constructive proof*.



Sylvester equation and Roth's removal rule

Matrices

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

are similar iff the Sylvester equation $A_{11}X - XA_{22} + A_{12} = 0$ is solvable.

Proof ("if" part): If $A_{11}X - XA_{22} + A_{12} = 0$ is solvable, then

$$\begin{aligned} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}^{-1} &= \begin{bmatrix} I & -X \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{11}X - XA_{22} + A_{12} \\ 0 & A_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix} \end{aligned}$$

This is what we call *constructive proof*.



Matrix Lyapunov equation / Matlab: `lyap(A,Q)`

Given $A \in \mathbb{R}^{n \times n}$ and $Q = Q' \in \mathbb{R}^{n \times n}$, find $X \in \mathbb{R}^{n \times n}$ such that

$$AX + XA' + Q = 0.$$

This is a particular case of Sylvester for $A_1 = A_2' = A$ and symmetric Q . As such, X exists for all Q and is unique iff $\lambda_i + \lambda_j \neq 0$, $\forall \lambda_i, \lambda_j \in \text{spec}(A)$.

If A is Hurwitz, i.e. $\text{spec}(A) \in \{z \in \mathbb{C} \mid \text{Re } z < 0\}$, then $\lambda_i + \lambda_j \neq 0$ and

$$X = \int_0^\infty e^{At} Q e^{A't} dt, \quad \forall Q.$$

Moreover, $X > 0$ whenever $Q > 0$ (and $X \geq 0$ whenever $Q \geq 0$).

Proof: Because $\|e^{At}\| = e^{\text{Re } \lambda_1 t} \rightarrow 0$ as $t \rightarrow \infty$,

we have $\|e^{At}\| \leq e^{-\alpha t}$ for some $\alpha > 0$ and $\|e^{A't}\| \leq e^{-\alpha t}$ for some $\alpha > 0$.

Thus, $\|e^{At} Q e^{A't}\| \leq \|Q\| e^{-2\alpha t}$ and $\int_0^\infty e^{-2\alpha t} dt < \infty$.

Moreover, $X = \int_0^\infty e^{At} Q e^{A't} dt$ is symmetric and $X \geq 0$.

Matrix Lyapunov equation / Matlab: `lyap(A,Q)`

Given $A \in \mathbb{R}^{n \times n}$ and $Q = Q' \in \mathbb{R}^{n \times n}$, find $X \in \mathbb{R}^{n \times n}$ such that

$$AX + XA' + Q = 0.$$

This is a particular case of Sylvester for $A_1 = A'_2 = A$ and symmetric Q . As such, X exists for all Q and is unique iff $\lambda_i + \lambda_j \neq 0$, $\forall \lambda_i, \lambda_j \in \text{spec}(A)$.

If A is Hurwitz, i.e. $\text{spec}(A) \in \{z \in \mathbb{C} \mid \text{Re } z < 0\}$, then $\lambda_i + \lambda_j \neq 0$ and

$$X = \int_0^\infty e^{At} Q e^{A't} dt, \quad \forall Q.$$

Moreover, $X > 0$ whenever $Q > 0$ (and $X \geq 0$ whenever $Q \geq 0$).

Proof: Because $\frac{d}{dt} e^{At} Q e^{A't} = A e^{At} Q e^{A't} + e^{At} Q e^{A't} A'$,

$$A \int_0^\infty e^{At} Q e^{A't} dt + \int_0^\infty e^{At} Q e^{A't} dt A = \int_0^\infty \frac{d}{dt} e^{At} Q e^{A't} dt = e^{At} Q e^{A't} \Big|_0^\infty = -Q$$

and the result follows by the fact that X is unique for every Hurwitz A . Now, if $Q > 0$ ($Q \geq 0$), then $e^{At} Q e^{A't} > 0$ ($e^{At} Q e^{A't} \geq 0$) for all t as well. Hence, so is X . $\square$

Matrix Stein equation / Matlab: `dlyap(A1,A2,Q)`

Given $A_1 \in \mathbb{R}^{p \times p}$, $A_2 \in \mathbb{R}^{m \times m}$, and $Q \in \mathbb{R}^{p \times m}$, find $X \in \mathbb{R}^{p \times m}$ such that

$$X = A_1 X A_2 + Q.$$

X exists for all Q and is unique iff $\lambda\mu \neq 1$, $\forall \lambda \in \text{spec}(A_1), \mu \in \text{spec}(A_2)$.

Example: Let $A_1 = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} a & 0 \\ 1 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$. In this case

$$\begin{aligned} - \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} &= \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} + \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} a & 0 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (a-1)x_{11} + x_{21} + x_{12} & x_{22} - x_{12} \\ (a+1)x_{21} + x_{22} & x_{22} \end{bmatrix} \end{aligned}$$

so $x_{22} = -q_{22}$, $x_{12} = q_{12} - q_{22}$,

$$(a+1)x_{21} = q_{22} - q_{21}, \quad \text{and} \quad (a-1)x_{11} = q_{22} - q_{12} - x_{21} - q_{11}.$$

Solvable $\forall Q$ iff $a \notin \{-1, 1\}$. Otherwise, depends on Q and is not unique.

Matrix Stein equation / Matlab: `dlyap(A1,A2,Q)`

Given $A_1 \in \mathbb{R}^{p \times p}$, $A_2 \in \mathbb{R}^{m \times m}$, and $Q \in \mathbb{R}^{p \times m}$, find $X \in \mathbb{R}^{p \times m}$ such that

$$X = A_1 X A_2 + Q.$$

X exists for all Q and is unique iff $\lambda\mu \neq 1$, $\forall \lambda \in \text{spec}(A_1), \mu \in \text{spec}(A_2)$.

Example: Let $A_1 = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} a & 0 \\ 1 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$. In this case

$$\begin{aligned} - \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} &= \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} + \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} a & 0 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} (a-1)x_{11} + x_{21} + x_{12} & x_{22} - x_{12} \\ (a+1)x_{21} + x_{22} & x_{22} \end{bmatrix} \end{aligned}$$

so $x_{22} = -q_{22}$, $x_{12} = q_{12} - q_{22}$,

$$(a+1)x_{21} = q_{22} - q_{21}, \quad \text{and} \quad (a-1)x_{11} = q_{22} - q_{12} - x_{21} - q_{11}.$$

Solvable $\forall Q$ iff $a \notin \{-1, 1\}$. Otherwise, depends on Q and is not unique.

Matrix discrete Lyapunov equation / Matlab: dlyap(A,Q)

Given $A \in \mathbb{R}^{n \times n}$ and $Q = Q' \in \mathbb{R}^{n \times n}$, find $X \in \mathbb{R}^{n \times n}$ such that

$$X = AXA' + Q.$$

This is a particular case of Stein for $A_1 = A_2' = A$ and symmetric Q . Hence, X exists for all Q and is unique iff $\lambda_i \lambda_j \neq 1$, $\forall \lambda_i, \lambda_j \in \text{spec}(A)$.

If A is Schur, i.e. $\text{spec}(A) \in \{z \in \mathbb{C} \mid |z| < 1\}$, then $\lambda_i \lambda_j \neq 1$ and

$$X = \sum_{i=0}^{\infty} A^i Q (A^i)'. \quad \forall Q.$$

Moreover, $X > 0$ whenever $Q > 0$ (and $X \geq 0$ whenever $Q \geq 0$).

Proof: Exercise (Stein, 1969).

Matrix discrete Lyapunov equation / Matlab: dlyap(A,Q)

Given $A \in \mathbb{R}^{n \times n}$ and $Q = Q' \in \mathbb{R}^{n \times n}$, find $X \in \mathbb{R}^{n \times n}$ such that

$$X = AXA' + Q.$$

This is a particular case of Stein for $A_1 = A_2' = A$ and symmetric Q . Hence, X exists for all Q and is unique iff $\lambda_i \lambda_j \neq 1$, $\forall \lambda_i, \lambda_j \in \text{spec}(A)$.

If A is Schur, i.e. $\text{spec}(A) \in \{z \in \mathbb{C} \mid |z| < 1\}$, then $\lambda_i \lambda_j \neq 1$ and

$$X = \sum_{t=0}^{\infty} A^t Q (A^t)', \quad \forall Q.$$

Moreover, $X > 0$ whenever $Q > 0$ (and $X \geq 0$ whenever $Q \geq 0$).

Proof: Because $\lim_{t \rightarrow \infty} A^t = 0$,

$$A \left(\sum_{t=0}^{\infty} A^t Q (A^t)' \right) A' + Q = \sum_{t=1}^{\infty} A^t Q (A^t)' + Q = \sum_{t=0}^{\infty} A^t Q (A^t)'$$

and the result follows by the fact that X is unique for every Schur A . Now, if $Q > 0$ ($Q \geq 0$), then $A^t Q (A^t)' > 0$ ($A^t Q (A^t)' \geq 0$) for all t as well. Hence, so is X . $\square$