12. Nonnegative matrices

- matrix norms and spectral radius
- linear dynamical systems
- Perron–Frobenius theorem
Norms

recall the three defining properties of a norm (of vectors or matrices)

• positive definiteness: \( \|x\| \geq 0 \) for all \( x \), and \( \|x\| = 0 \) only if \( x = 0 \)

• homogeneity: \( \|\gamma x\| = |\gamma| \|x\| \) for all \( x \) and all scalars \( \gamma \)

• triangle inequality: \( \|x + y\| \leq \|x\| + \|y\| \) for all \( x, y \)

Examples of vector norms (on \( \mathbb{R}^n \) or \( \mathbb{C}^n \))

• Euclidean norm \( \|x\|_2 = (|x_1|^2 + \cdots + |x_n|^2)^{1/2} \) (denoted by \( \|x\| \) in this course)

• Chebyshev norm (infinity-norm): \( \|x\|_\infty = \max\{|x_1|, |x_2|, \ldots, |x_n|\} \)

• 1-norm: \( \|x\|_1 = |x_1| + |x_2| + \cdots + |x_n| \)

Examples of matrix norms (on \( \mathbb{R}^{m\times n} \) or \( \mathbb{C}^{m\times n} \))

• Frobenius norm \( \|X\|_F = \sum_{i=1}^{m} \sum_{j=1}^{n} |X_{ij}|^2 \)

• matrix 2-norm (spectral norm) \( \|X\|_2 = \sigma_1(X) \) (maximum singular value)

• max-row-sum norm \( \|X\|_\infty = \max \sum_{j=1}^{n} |X_{ij}| \)

Nonnegative matrices 12.2
Submultipliciative matrix norms

for matrix norms on $\mathbb{R}^{n \times n}$ or $\mathbb{C}^{n \times n}$ one often requires a fourth property:

$$\|AB\| \leq \|A\| \|B\| \quad \text{for all } A, B$$

- such a norm is called submultiplicative
- for a submultiplicative norm, $\|A^k\| \leq \|A\|^k$ for positive integer $k$

Examples: the following matrix norms are submultiplicative

- Frobenius norm
- matrix 2-norm
- max-row-sum norm

Exercise: show that the following matrix norm is not submultiplicative

$$\|X\| = \max_{i=1,\ldots,n} \max_{j=1,\ldots,n} |X_{ij}|$$
Spectral radius

the *spectral radius* of an \( n \times n \) matrix \( A \) is defined as

\[
\rho(A) = \max_{i=1,...,n} |\lambda_i(A)|
\]

Example

\[
A = \begin{bmatrix}
-1 & -2 & 6 \\
-2 & 1 & 5 \\
2 & -5 & -4
\end{bmatrix}
\]

- eigenvalues of \( A \):

\[
\lambda_1 = -2.72 + 2.91j, \quad \lambda_2 = -2.72 - 2.91j, \quad \lambda_3 = 1.45
\]

- spectral radius is \( \rho(A) = |\lambda_1| = |\lambda_2| = 3.99 \)
Spectral radius of normal matrix

for a normal matrix,

\[ \rho(A) = \|A\|_2 \]

- recall definition of normal matrix on page 10.12: \( AA^H = A^H A \)
- normal matrices include symmetric, skew-symmetric, orthogonal matrices
- Schur decomposition of normal matrix is of the form

\[ A = UDU^H \quad \text{with } U \text{ unitary, } D \text{ diagonal} \]

diagonal entries of \( D \) are the eigenvalues \( \lambda_i \) of \( A \)
- therefore, singular values \( \sigma_k \) are the absolute values of the eigenvalues, so

\[ \rho(A) = \max_{i=1,\ldots,n} |\lambda_i| = \max_{k=1,\ldots,n} \sigma_k = \|A\|_2 \]

for non-normal matrices \( \rho(A) \) and \( \|A\|_2 \) can be very different: for \( A \) on page 12.4,

\[ \rho(A) = 3.99, \quad \|A\|_2 = 9.41 \]
Spectral radius is not a matrix norm

the spectral radius is not a norm on $\mathbb{R}^{n \times n}$ or $\mathbb{C}^{n \times n}$

• nonzero matrix can have zero spectral radius (all eigenvalues are zero)

\[
A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}
\]

• triangle inequality does not hold for spectral radius

\[
A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 \\ 3 & -1 \end{bmatrix}, \quad A + B = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}
\]

here, $\rho(A) = \rho(B) = 1$, but

\[
\rho(A + B) = 3 > \rho(A) + \rho(B)
\]
Norm bound on spectral radius

for any submultiplicative matrix norm,

$$\rho(A) \leq \|A\|$$  \hspace{1cm} (1)

Proof

- choose a vector norm that is consistent with the matrix norm, \textit{i.e.},

$$\|Ax\| \leq \|A\|\|x\| \quad \text{for all } A \text{ and all } x$$

- for example, define vector norm as \(\|x\| = \|xy^T\|\) where \(y \neq 0\) is a fixed \(n\)-vector:

$$\|Ax\| = \|Axy^T\| \leq \|A\|\|xy^T\| = \|A\|\|x\|$$

inequality follows from submultiplicative property of the matrix norm

- now, if \(x\) is eigenvector with eigenvalue \(\lambda\) and \(|\lambda| = \rho(A)\), then (1) follows from:

$$\rho(A)\|x\| = |\lambda|\|x\| = \|\lambda x\| = \|Ax\| \leq \|A\|\|x\|$$
Outline

- matrix norms and spectral radius
- **linear dynamical systems**
- Perron–Frobenius theorem
Discrete-time linear dynamical system

\[ x(t + 1) = A(t)x(t), \quad t = 0, 1, 2, \ldots \]

- \( t \) denotes time or period
- \( n \)-vector \( x(t) \) is the state at time \( t \)
- \( n \times n \) matrix \( A(t) \) is the dynamics matrix at time \( t \)
- current state is a linear function of the previous state
- system is time-invariant if \( A(t) = A \) does not depend on time

for a time-invariant system,

\[ x(t) = A^t x(0), \quad t = 0, 1, 2, \ldots \]

- properties of matrix power \( A^t \) determine dynamical behavior
- in this lecture we will be interested in systems with \( A_{ij} \geq 0 \) for all \( i, j \)
Autoregressive model

\[ x(t + 1) = A_1 x(t) + A_2 x(t - 1) + \cdots + A_p x(t - p + 1) \]

- state evolution is described by a homogeneous difference equation
- current state is a linear function of the \( p \) previous states
- linear dynamical system \( x(t + 1) = A x(t) \) is a special case with \( p = 1 \)
- for \( p > 1 \), can be written as a linear dynamical system \( \tilde{x}(t + 1) = \tilde{A} \tilde{x}(t) \):

\[
\begin{bmatrix}
    x(t + 1) \\
    x(t) \\
    x(t - 1) \\
    \vdots \\
    x(t - p + 3) \\
    x(t - p + 2)
\end{bmatrix}
= \begin{bmatrix}
    A_1 & A_2 & \cdots & A_{p-1} & A_p \\
    I & 0 & \cdots & 0 & 0 \\
    0 & I & \cdots & 0 & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & I & 0
\end{bmatrix}
\begin{bmatrix}
    x(t) \\
    x(t - 1) \\
    x(t - 2) \\
    \vdots \\
    x(t - p + 2) \\
    x(t - p + 1)
\end{bmatrix}
\]
Population dynamics

• $x(t) \in \mathbb{R}^{100}$ gives age distribution of population in year $t$ (say, on January 1)

• $x_i(t)$ for $i < 100$ is the number of people with age $i - 1$

• $x_{100}(t)$ is the number of people aged 99 and above

• we exclude changes due to immigration

Linear dynamical model

$$
\begin{bmatrix}
    x_1(t+1) \\
    x_2(t+1) \\
    x_3(t+1) \\
    \vdots \\
    x_{100}(t+1)
\end{bmatrix}
= 
\begin{bmatrix}
    b_1 & b_2 & \cdots & b_{99} & b_{100} \\
    1 - d_1 & 0 & \cdots & 0 & 0 \\
    0 & 1 - d_2 & \cdots & 0 & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & 1 - d_{99} & 1 - d_{100}
\end{bmatrix}
\begin{bmatrix}
    x_1(t) \\
    x_2(t) \\
    x_3(t) \\
    \vdots \\
    x_{100}(t)
\end{bmatrix}
$$

• $b_i \geq 0$ is number of births per person with age $i - 1$

• $0 < d_i < 1$ is number of deaths per person with age $i - 1$
Epidemic dynamics

4-vector $x(t)$ gives proportion of population in 4 infection states

- susceptible: can acquire the disease the next day
- infected: have the disease
- recovered: had the disease, recovered, now immune
- deceased: had the disease, and unfortunately died

Example

$$
\begin{bmatrix}
    x_s(t+1) \\
    x_i(t+1) \\
    x_r(t+1) \\
    x_d(t+1)
\end{bmatrix} =
\begin{bmatrix}
    0.95 & 0.04 & 0 & 0 \\
    0.05 & 0.85 & 0 & 0 \\
    0 & 0.10 & 1 & 0 \\
    0 & 0.01 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    x_s(t) \\
    x_i(t) \\
    x_r(t) \\
    x_d(t)
\end{bmatrix}
$$

each day,

- 5% of susceptible population acquires the disease, 95% remains susceptible
- 4% of infected recover without immunity, 10% with immunity, 1% die
Simulation of epidemic dynamics

- figure shows $x(t)$ starting at $x(1) = (1, 0, 0, 0)$
- $x(t) \rightarrow (0.0, 0.0, 0.91, 0.09)$ (an eigenvector of $A$ with eigenvalue 1)
Finite Markov chain

- \( s(t) \) is a random variable, with possible values 1, 2, \ldots, \( n \) (the \( n \) “states”)
- \( x_i(t) \) is the probability that \( s(t) = i \)
- \( A_{ij} \) is probability that \( s(t + 1) = i \), given that \( s(t) = j \):

\[
x_i(t + 1) = A_{i1}x_1(t) + A_{i2}x_2(t) + \cdots + A_{in}x_n(t)
\]

hence, \( A \) is nonnegative, with \( A_{1j} + A_{2j} + \cdots + A_{nj} = 1 \) in each column

- matrix \( A^T \) is called the state transition matrix

**Example**

\[
A = \begin{bmatrix}
0.8 & 0.4 & 0.9 \\
0.2 & 0.1 & 0 \\
0 & 0.5 & 0.1
\end{bmatrix}
\]
Eigenvalues of matrix power

let $A$ be a square matrix with Schur decomposition

$$A = UTU^H$$

diagonal elements of $T$ are eigenvalues $\lambda_1, \ldots, \lambda_n$ of $A$

- if $k$ is a positive integer, then $A^k = UT^kU^H$
- this is a Schur decomposition: $T^k$ is upper triangular with diagonal elements $T_{11}^k, T_{22}^k, \ldots, T_{nn}^k$

- eigenvalues of $A^k$ are $\lambda_1^k, \lambda_2^k, \ldots, \lambda_n^k$
- spectral radius is $\rho(A^k) = \rho(A)^k$
- one can also show that

$$\lim_{k \to \infty} A^k = 0 \iff \rho(A) < 1$$
Outline

- matrix norms and spectral radius
- linear dynamical systems
- Perron–Frobenius theorem
Perron–Frobenius theorem

suppose $A \in \mathbb{R}^{n \times n}$ is nonnegative (componentwise) and irreducible

Perron root

- $\rho(A)$ is an eigenvalue of $A$, called the Perron root
- $\rho(A) > 0$
- the eigenvalue $\rho(A)$ has algebraic (hence, geometric) multiplicity one

Perron vector

- $A$ has a positive eigenvector with eigenvalue $\rho(A)$
- normalized to satisfy $\mathbf{1}^T x = 1$ this vector is unique, and called the Perron vector:

$$A x = \rho(A) x, \quad x > 0, \quad \mathbf{1}^T x = 1$$

- $A$ has no other nonnegative eigenvectors (except multiples of the Perron vector)

Other eigenvalues: if $A$ is positive, eigenvalues $\lambda \neq \rho(A)$ satisfy $|\lambda| < \rho(A)$
Nonnegative matrices

some weaker results hold if $A$ is nonnegative, but not irreducible

- $\rho(A)$ is an eigenvalue of $A$
- $A$ has a nonnegative eigenvector with eigenvalue $\rho(A)$
consider a finite Markov chain (page 12.13), and assume $A$ is irreducible

- the state transition matrix $A^T$ is also irreducible
- since $\sum_i A_{ij} = 1$ for all $j$, the vector $\mathbf{1}$ is an eigenvector of $A^T$ with eigenvalue $1$:

$$A^T \mathbf{1} = \mathbf{1}$$

- from the definition of $\rho(A)$ and the norm inequality on page 12.7

$$1 \leq \rho(A) = \rho(A^T) \leq \|A^T\|_\infty = 1$$

hence, $\rho(A) = 1$

- there is a unique positive stationary probability vector (the Perron vector of $A$):

$$Ax = x, \quad x > 0, \quad \mathbf{1}^T x = 1$$

for the example on page 12.13, the stationary probability distribution is

$$x = (0.743, 0.165, 0.092)$$