2. LINEAR ALGEBRA

Outline

1. Definitions
2. Linear least squares problem
3. QR factorization
4. Singular value decomposition (SVD)
5. Pseudo-inverse
6. Eigenvalue decomposition (EVD)

Definitions

Vector norm

- Let $\mathbf{x} \in \mathbb{C}^N$ be an N-dimensional complex vector.
- The Euclidean norm (2-norm) of $\mathbf{x}$ is

$$\|\mathbf{x}\| := \left(\sum_{i=1}^{N} |x_i|^2 \right)^{1/2} = \left(\sum_{i=1}^{N} \bar{x}_i x_i \right)^{1/2} = (\mathbf{x}^{\mathrm{H}} \mathbf{x})^{1/2}$$

Matrix norms

- Let $\mathbf{A} \in \mathbb{C}^{M \times N}$ be an $M \times N$ complex matrix.
- The *induced matrix 2-norm* (spectral norm, operator norm) is

$$\|\mathbf{A}\| := \max_{\mathbf{x}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} \qquad \text{or} \qquad \|\mathbf{A}\|^2 = \max_{\mathbf{x}} \frac{\mathbf{x}^{\mathrm{H}} \mathbf{A}^{\mathrm{H}} \mathbf{A} \mathbf{x}}{\mathbf{x}^{\mathrm{H}} \mathbf{x}}$$

- The *Frobenius norm* of $\mathbf{A}$ is

$$\|\mathbf{A}\|_{\mathrm{F}} = \left(\sum_{i=1}^{M} \sum_{j=1}^{N} |a_{ij}|^2 \right)^{1/2}$$

Definitions

Linear independence

- A collection of vectors $\{\mathbf{x}_i\}$ is called linear independent if

$$\alpha_1 \mathbf{x}_1 + \cdots + \alpha_N \mathbf{x}_N = 0 \quad \Leftrightarrow \quad \alpha_1 = \cdots = \alpha_N = 0\,.$$

Subspaces

- The space $\mathcal{H}$ spanned by a collection of vectors $\{\mathbf{x}_i\}$

$$\mathcal{H} := \{\alpha_1 \mathbf{x}_1 + \cdots + \alpha_N \mathbf{x}_N \mid \alpha_i \in \mathbb{C}\,,\ \forall i\}$$

  is called a *linear subspace*

- Example subspaces:

Range (column span) of **A**:	$\text{ran}(\mathbf{A}) = \{\mathbf{A}\mathbf{x}:\ \mathbf{x} \in \mathbb{C}^N\}$
Kernel (row nullspace) of **A**:	$\ker(\mathbf{A}) = \{\mathbf{x} \in \mathbb{C}^N:\ \mathbf{A}\mathbf{x} = 0\}$

Definitions

Basis

- An independent collection of vectors $\{\mathbf{x}_i\}$ that together span a subspace is called a *basis* for that subspace.
- If the vectors are orthogonal ($\mathbf{x}_i^{\mathrm{H}}\mathbf{x}_j = 0,\ i \neq j$) ➠ *orthogonal basis*.
- If the vectors are orthonormal ($\mathbf{x}_i^{\mathrm{H}}\mathbf{x}_j = 0,\ i \neq j$ and $\|\mathbf{x}_i\| = 1$) ➠ *orthonormal basis*.

Rank

- The *rank* of a matrix **A** is the max. nr. of independent columns (or rows) of **A**.

Prototype rank-1 matrix: $\mathbf{A} = \mathbf{a}\mathbf{b}^{\mathrm{H}}$

Prototype rank-2 matrix: $\mathbf{A} = \mathbf{a}\mathbf{b}^{\mathrm{H}} + \mathbf{c}\mathbf{d}^{\mathrm{H}}$

Definitions

Unitary matrix

- A square matrix $\mathbf{U}$ is called *unitary* if $\mathbf{U}^{\mathrm{H}}\mathbf{U} = \mathbf{I}$ and $\mathbf{U}\mathbf{U}^{\mathrm{H}} = \mathbf{I}$.
- Properties:
  - A unitary matrix looks like a rotation and/or a reflection.
  - Its norm is $\|\mathbf{U}\| = 1$.
  - Its columns and rows are orthonormal.

Isometry

- A tall rectangular matrix $\hat{\mathbf{U}}$ is called an isometry if $\hat{\mathbf{U}}^{\mathrm{H}}\hat{\mathbf{U}} = \mathbf{I}$.
  - Its columns are orthonormal basis of a subspace (not the complete space).
  - Its norm is $\|\hat{\mathbf{U}}\| = 1$.
  - There is an orthogonal complement $\hat{\mathbf{U}}^{\perp}$ of $\hat{\mathbf{U}}$ such that $\mathbf{U} = [\hat{\mathbf{U}} \;\; \hat{\mathbf{U}}^{\perp}]$ is unitary.

Definitions

Projection

- A square matrix $\mathbf{P}$ is a projection if $\mathbf{PP} = \mathbf{P}$.
- It is an orthogonal projection if also $\mathbf{P}^{\mathrm{H}} = \mathbf{P}$.
  - The norm of an orthogonal projection is $\|\mathbf{P}\| = 1$.
  - For an isometry $\hat{\mathbf{U}}$, the matrix $\mathbf{P} = \hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{H}}$ is an orthogonal projection (onto the space spanned by the columns of $\hat{\mathbf{U}}$). This is the general form of a projection.
- Suppose $\mathbf{U} = [\underbrace{\hat{\mathbf{U}}}_{d} \quad \underbrace{\hat{\mathbf{U}}^{\perp}}_{M-d}]$ is unitary. Then, from $\mathbf{UU}^{\mathrm{H}} = \mathbf{I}_M$:

$$\hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{H}} + \hat{\mathbf{U}}^{\perp}(\hat{\mathbf{U}}^{\perp})^{\mathrm{H}} = \mathbf{I}_M\,, \qquad \hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{H}} = \mathbf{P}\,, \qquad \hat{\mathbf{U}}^{\perp}(\hat{\mathbf{U}}^{\perp})^{\mathrm{H}} = \mathbf{P}^{\perp} \;=\; \mathbf{I}_M - \mathbf{P}$$

- Any vector $\mathbf{x} \in \mathbb{C}^M$ can be decomposed into $\mathbf{x} = \hat{\mathbf{x}} + \hat{\mathbf{x}}^{\perp}$, where $\hat{\mathbf{x}} \perp \hat{\mathbf{x}}^{\perp}$,

$$\hat{\mathbf{x}} = \mathbf{Px} \in \mathrm{ran}(\hat{\mathbf{U}})\,, \qquad \hat{\mathbf{x}}^{\perp} = \mathbf{P}^{\perp}\mathbf{x} \in \mathrm{ran}(\hat{\mathbf{U}}^{\perp})$$

Definitions

Projection onto the column span of A

- Suppose $\mathbf{A}$ is tall and $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ is invertible. Then

$$\mathbf{P}_{\mathbf{A}} := \mathbf{A}(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}, \qquad \mathbf{P}_{\mathbf{A}}^{\perp} := \mathbf{I} - \mathbf{A}(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}$$

  are orthogonal projections, onto the range of $\mathbf{A}$ and kernel of $\mathbf{A}^{\mathrm{H}}$, resp.

- Proof:

  Verify that $\mathbf{PP} = \mathbf{P}$ and $\mathbf{P}^{\mathrm{H}} = \mathbf{P}$, hence $\mathbf{P}$ is an orthogonal projection.

  If $\mathbf{b} \in \mathrm{ran}(\mathbf{A})$, then $\mathbf{b} = \mathbf{Ax}$ for some $\mathbf{x}$.

  Hence

$$\mathbf{P}_{\mathbf{A}}\mathbf{b} = \mathbf{A}(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{Ax} = \mathbf{b}$$

  so that $\mathbf{b}$ is invariant under $\mathbf{P}_{\mathbf{A}}$.

  If $\mathbf{b} \perp \mathrm{ran}(\mathbf{A})$, then $\mathbf{b} \in \ker(\mathbf{A}^{\mathrm{H}})$, or $\mathbf{A}^{\mathrm{H}}\mathbf{b} = \mathbf{0}$. Hence $\mathbf{P}_{\mathbf{A}}\mathbf{b} = \mathbf{0}$.

Linear least squares problem

- Given $\mathbf{A}$, $\mathbf{b}$, find

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \| \mathbf{A}\mathbf{x} - \mathbf{b} \|^2$$

- Solution:

  Write $\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2$, where $\mathbf{b}_1 \in \text{ran}(\mathbf{A})$, $\mathbf{b}_2 \perp \text{ran}(\mathbf{A})$.

  Then

$$\mathbf{b}_1 = \mathbf{P}_{\mathbf{A}}\mathbf{b} = \mathbf{A}(\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{b}$$

$$\mathbf{A}\mathbf{x} - \mathbf{b} = \mathbf{A}\left\{\mathbf{x} - (\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{b}\right\} - \mathbf{b}_2$$

  Note that the two terms are orthogonal. Thus

$$\| \mathbf{A}\mathbf{x} - \mathbf{b} \|^2 = \| \mathbf{A}\left\{\mathbf{x} - (\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{b}\right\} \|^2 + \| \mathbf{b}_2 \|^2$$

  To minimize the error, set $\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{b}$.

QR factorization

- Let **A** be an $N \times N$ square full rank matrix.

  Then there is a decomposition

$$\mathbf{A} = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_N\end{bmatrix} = \begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_N\end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ 0 & r_{22} & \cdots & r_{2N} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & r_{NN} \end{bmatrix} = \mathbf{QR}$$

  Here, **Q** is a unitary matrix, **R** is upper triangular and square.

- Interpretation:
  - $\mathbf{q}_1$ is a normalized vector with the same direction as $\mathbf{a}_1$.
  - $[\mathbf{q}_1 \ \mathbf{q}_2]$ is an isometry spanning the same space as $[\mathbf{a}_1 \ \mathbf{a}_2]$.
  - $[\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3]$ is an isometry spanning the same space as $[\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3]$.
  - Etc.

QR factorization

- Let $\mathbf{A}$ be an $M \times N$ tall ($M \geq N$) matrix.

  Then there is a decomposition

$$\mathbf{A} = \mathbf{QR} = [\hat{\mathbf{Q}} \quad \hat{\mathbf{Q}}^{\perp}] \begin{bmatrix} \hat{\mathbf{R}} \\ \mathbf{0} \end{bmatrix} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$$

  Here, $\mathbf{Q}$ is a unitary matrix, $\hat{\mathbf{R}}$ is upper triangular and square.

- Properties:

  - $\mathbf{R}$ is upper triangular with $M - N$ zero rows added.
  - $\mathbf{A} = \hat{\mathbf{Q}}\hat{\mathbf{R}}$ is an "economy-size" QR-decomposition.
  - If $\hat{\mathbf{R}}$ is full rank, the columns of $\hat{\mathbf{Q}}$ span the range of $\mathbf{A}$.
  - If $\hat{\mathbf{R}}$ is not full rank, the column span of $\hat{\mathbf{Q}}$ is too large.

Singular value decomposition

- For any matrix $\mathbf{X}$, there is a decomposition

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{H}}$$

Here, $\mathbf{U}$ and $\mathbf{V}$ are unitary, and $\mathbf{\Sigma}$ is diagonal, with positive real entries.

- Properties:
  - The columns $\mathbf{u}_i$ of $\mathbf{U}$ are called the left singular vectors.
  - The columns $\mathbf{v}_i$ of $\mathbf{V}$ are called the right singular vectors.
  - The diagonal entries σ_i of $\mathbf{\Sigma}$ are called the singular values.
  - They are positive and real, and usually sorted such that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$$

Singular value decomposition

- More specifically, for an $M \times N$ tall ($M \geq N$) matrix **X**:

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{H}} = [\hat{\mathbf{U}} \quad \hat{\mathbf{U}}^{\perp}] \left[\begin{array}{cc|cc} \sigma_1 & & & \\ & \sigma_d & & \\ \hline & & 0 & \\ & & & 0 \\ \hline 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \end{array}\right] \begin{bmatrix} \hat{\mathbf{V}}^{\mathrm{H}} \\ (\hat{\mathbf{V}}^{\perp})^{\mathrm{H}} \end{bmatrix}$$

$$\mathbf{U} : M \times M, \quad \mathbf{\Sigma} : M \times N, \quad \mathbf{V} : N \times N$$

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_d > \sigma_{d+1} = \cdots = \sigma_N = 0$$

- *'Economy size' SVD:* $\mathbf{X} = \hat{\mathbf{U}}\hat{\Sigma}\hat{\mathbf{V}}^{\mathrm{H}}$, where $\hat{\mathbf{\Sigma}} : d \times d$, containing $\sigma_1, \cdots, \sigma_d$.

Singular value decomposition

Some SVD facts

- The rank of **X** is d, the number of nonzero singular values.

- $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{H}} \quad \Leftrightarrow \quad \mathbf{X}^{\mathrm{H}} = \mathbf{V}\Sigma\mathbf{U}^{\mathrm{H}} \quad \Leftrightarrow \quad \mathbf{X}\mathbf{V} = \mathbf{U}\Sigma \quad \Leftrightarrow \quad \mathbf{X}^{\mathrm{H}}\mathbf{U} = \mathbf{V}\Sigma$

  $\Rightarrow$ The columns of $\hat{\mathbf{U}}$ ($\hat{\mathbf{U}}^{\perp}$) are orthonormal basis for range of **X** (kernel of $\mathbf{X}^{\mathrm{H}}$).

  $\Rightarrow$ The columns of $\hat{\mathbf{V}}$ ($\hat{\mathbf{V}}^{\perp}$) are orthonormal basis for range of $\mathbf{X}^{\mathrm{H}}$ (kernel of **X**).

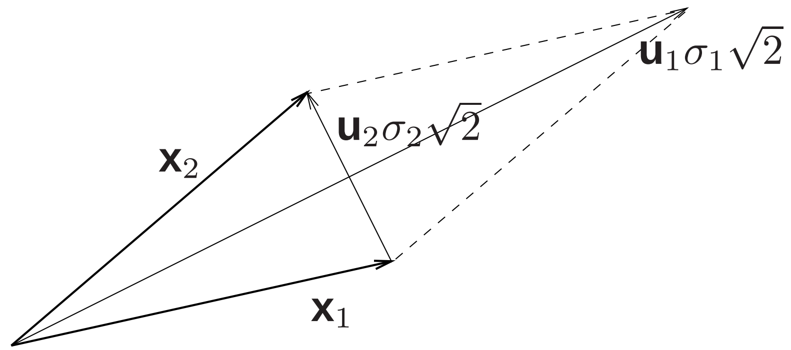
- The norm of **X** or $\mathbf{X}^{\mathrm{H}}$ is $\|\mathbf{X}\| = \|\mathbf{X}^{\mathrm{H}}\| = \sigma_1$, the largest singular value.

  The norm is attained on the corresponding singular vectors $\mathbf{u}_1$ and $\mathbf{v}_1$:

$$\mathbf{X}\mathbf{v}_1 = \mathbf{u}_1\sigma_1 \qquad \mathbf{X}^{\mathrm{H}}\mathbf{u}_1 = \mathbf{v}_1\sigma_1$$

Singular value decomposition

Geometrical interpretation



Construction of the left singular vectors and values of the matrix $\mathbf{X} = [\mathbf{x}_1 \ \ \mathbf{x}_2]$, where $\mathbf{x}_1$ and $\mathbf{x}_2$ have *equal length*.

- The largest singular vector $\mathbf{u}_1$ is in the direction of the sum of $\mathbf{x}_1$ and $\mathbf{x}_2$: the 'common' direction of the two vectors.

  Singular value: $\sigma_1 = \|\mathbf{x}_1 + \mathbf{x}_2\|/\sqrt{2}$.

- The smallest singular vector $\mathbf{u}_2$ depends on the difference $\mathbf{x}_2 - \mathbf{x}_1$.

  Singular value: $\sigma_2 = \|\mathbf{x}_2 - \mathbf{x}_1\|/\sqrt{2}$.

Singular value decomposition

Connections between the SVD and QR factorizations

- The QR factorization of a tall ($M \geq N$) matrix **X** is

$$\mathbf{X} = \mathbf{QR} = [\hat{\mathbf{Q}} \quad \hat{\mathbf{Q}}^{\perp}] \begin{bmatrix} \hat{\mathbf{R}} \\ 0 \end{bmatrix}$$

- The QR factorization can be used as a starting point for the SVD of **X**:

  First compute the SVD of $\hat{\mathbf{R}}$

$$\hat{\mathbf{R}} = \hat{\mathbf{U}}_R \hat{\Sigma}_R \hat{\mathbf{V}}_R^{\mathrm{H}}$$

  so that the SVD of **X** is

$$\mathbf{X} = (\hat{\mathbf{Q}} \hat{\mathbf{U}}_R) \hat{\Sigma}_R \hat{\mathbf{V}}_R^{\mathrm{H}}$$

  X and **R** have the same Σ and **V**.

Pseudo-inverse

Full rank pseudo-inverse

- $\mathbf{X} : M \times N$, tall ($M \geq N$), full rank.

  The *pseudo-inverse* of $\mathbf{X}$ is $\mathbf{X}^\dagger = (\mathbf{X}^{\mathrm{H}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{H}}$.

- It satisfies $\mathbf{X}^\dagger\mathbf{X} = \mathbf{I}_N$ (i.e., $\mathbf{X}^\dagger$ is an inverse on the “short space”).

- Also, $\mathbf{X}\mathbf{X}^\dagger = \mathbf{P}$: a projection onto the column span of $\mathbf{X}$.

Rank-deficient pseudo-inverse

- $\mathbf{X} : M \times N$, tall ($M \geq N$), rank-d, with ‘economy size’ SVD $\mathbf{X} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^{\mathrm{H}}$.

  The pseudo-inverse of $\mathbf{X}$ is $\mathbf{X}^\dagger = \hat{\mathbf{V}}\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{U}}^{\mathrm{H}}$.

- It satisfies $\mathbf{X}\mathbf{X}^\dagger = \hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathrm{H}} = \mathbf{P}_c\,, \quad \mathbf{X}^\dagger\mathbf{X} = \hat{\mathbf{V}}\hat{\mathbf{V}}^{\mathrm{H}} = \mathbf{P}_r$.
  - The norm of $\mathbf{X}^\dagger$ is $\|\mathbf{X}^\dagger\| = \sigma_d^{-1}$.
  - The *condition number* of $\mathbf{X}$ is $c(\mathbf{X}) := \frac{\sigma_1}{\sigma_d}$.

    If it is large, then $\mathbf{X}$ is hard to invert ($\mathbf{X}^\dagger$ is sensitive to small changes).

Pseudo-inverse

Interpretation of condition number

- The condition number gives the relative sensitivity of the solution of linear systems of equations.

- Illustration:

$$\begin{aligned} \mathbf{Ax} = \mathbf{b} \quad &\Rightarrow \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \\ \mathbf{b}^1 = \mathbf{b} + \mathbf{e} \quad &\Rightarrow \quad \mathbf{x}^1 = \mathbf{x} + \mathbf{A}^{-1}\mathbf{e} \end{aligned}$$

Define $\sigma_1 = \|\mathbf{A}\|$, $\quad \sigma_N^{-1} = \|\mathbf{A}^{-1}\|$. $\quad$ Use $\|\mathbf{Ax}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|$.

Then

$$\begin{aligned} \|\mathbf{A}^{-1}\mathbf{e}\| &\leq \sigma_N^{-1}\|\mathbf{e}\| \\ \|\mathbf{b}\| &\leq \sigma_1\|\mathbf{x}\| \\ \frac{\|\mathbf{x}^1 - \mathbf{x}\|}{\|\mathbf{x}\|} &\leq \sigma_N^{-1}\frac{\|\mathbf{e}\|}{\|\mathbf{x}\|} \leq \sigma_N^{-1}\sigma_1\frac{\|\mathbf{e}\|}{\|\mathbf{b}\|} \end{aligned}$$

Pseudo-inverse

Rank approximation

- $\mathbf{X} : M \times N$, with SVD $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathrm{H}}$.

- To improve the condition number of **X**, we can set the small σ_i equal to zero.

  This leads to a low rank approximation of **X**.

- Illustration:

  - Choose a threshold ϵ, and suppose d singular values are larger than ϵ.
  - $\hat{\mathbf{U}}$: first d columns of **U**, $\hat{\mathbf{V}}$: first d columns of **V**, $\hat{\mathbf{\Sigma}}$: top-left $d \times d$ block of $\mathbf{\Sigma}$.
  - Then $\hat{\mathbf{X}} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^{\mathrm{H}}$ is a rank-d approximant of **X**, with error

$$\begin{aligned} \|\mathbf{X} - \hat{\mathbf{X}}\| &= \sigma_{d+1} \\ \|\mathbf{X} - \hat{\mathbf{X}}\|_{\mathrm{F}}^2 &= \sigma_{d+1}^2 + \cdots + \sigma_N^2 \end{aligned}$$

Eigenvalue decomposition

Definition

- The **eigenvalue problem** is $\mathbf{Ax} = \lambda \mathbf{x} \quad \Leftrightarrow \quad (\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = 0$.
- Any λ that makes $\mathbf{A} - \lambda \mathbf{I}$ singular is called an eigenvalue
- The corresponding $\mathbf{x}$ is the eigenvector (invariant vector).
- Stacking the results gives

$$
\begin{aligned}
\mathbf{A}[\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots] &= [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots] \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{bmatrix} \\
\Leftrightarrow \qquad \mathbf{AT} &= \mathbf{T\Lambda}
\end{aligned}
$$

- A "regular" matrix **A** has an **eigenvalue decomposition:**

$$\mathbf{A} = \mathbf{T\Lambda T}^{-1}, \qquad \text{where } \mathbf{T} \text{ is invertible and } \mathbf{\Lambda} \text{ is diagonal.}$$

This decomposition might not exist if eigenvalues are repeated.

Eigenvalue decomposition

Schur decomposition

- Suppose $\mathbf{T}$ has QR factorization $\mathbf{T} = \mathbf{QR}_T \quad \Rightarrow \quad \mathbf{T}^{-1} = \mathbf{R}_T^{-1}\mathbf{Q}^{\mathrm{H}}$. Hence

$$\mathbf{A} \;=\; \mathbf{QR}_T \mathbf{\Lambda} \mathbf{R}_T^{-1}\mathbf{Q}^{\mathrm{H}} \;=\; \mathbf{QRQ}^{\mathrm{H}}$$

- $\mathbf{A} = \mathbf{QRQ}^{\mathrm{H}}$, with $\mathbf{Q}$ unitary and $\mathbf{R}$ upper triangular, is a **Schur decomposition**.
- Properties:
  - $\mathbf{R}$ has the eigenvalues of $\mathbf{A}$ on the diagonal.
  - This decomposition always exists.
  - $\mathbf{Q}$ gives information about “eigen-subspaces” (invariant subspaces).

    But $\mathbf{Q}$ does not contain the eigenvectors.

Eigenvalue decomposition

Connection of the eigenvalue decomposition with the SVD

- Starting from the SVD we obtain

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{H}} \quad \Rightarrow \quad \mathbf{X}\mathbf{X}^{\mathrm{H}} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{H}}\mathbf{V}\Sigma\mathbf{U}^{\mathrm{H}} = \mathbf{U}\Sigma^2\mathbf{U}^{\mathrm{H}} = \mathbf{U}\Lambda\mathbf{U}^{\mathrm{H}}$$

- Hence, we can state
  - The eigenvalues of $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ are the singular values of $\mathbf{X}$, squared
    - ➠ The eigenvalues of $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ are real
  - The eigenvectors of $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ are the left singular vectors of $\mathbf{X}$
    - ➠ $\mathbf{U}$ is a unitary matrix
  - The SVD always exists
    - ➠ The eigenvalue decomposition of $\mathbf{X}\mathbf{X}^{\mathrm{H}}$ always exists

Eigenvalue decomposition

Noise covariance matrix

- Suppose we have M antennas, and receive only noise:

$$\mathbf{e}(k) = \begin{bmatrix} e_1(k) \\ \vdots \\ e_M(k) \end{bmatrix} = \mathbf{e}_k$$

- Collect the samples in a matrix $\mathbf{E} = [\mathbf{e}_1 \quad \mathbf{e}_2 \cdots \mathbf{e}_N] : \quad M \times N$

- The noise covariance matrix is

$$\mathbf{R}_e := \mathrm{E}(\mathbf{e}\mathbf{e}^{\mathrm{H}}) \simeq \hat{\mathbf{R}}_e := \frac{1}{N}\sum \mathbf{e}_k\mathbf{e}_k^{\mathrm{H}} = \frac{1}{N}\mathbf{E}\mathbf{E}^{\mathrm{H}}$$

  - $\mathbf{R}_e$ is hermitian: $\mathbf{R}_e^{\mathrm{H}} = \mathbf{R}_e$.
  - If noise is independent among sensors (*spatially white*), then $\mathbf{R}_e$ is diagonal.
  - If noise is independent identically distributed (i.i.d.), then $\mathbf{R}_e = \sigma^2\mathbf{I}$.
  - Hence, all eigenvalues of $\mathbf{R}_e$ are equal to σ^2 (the noise power).

Eigenvalue decomposition

EVD of a data matrix

- Suppose we collect a data matrix $\mathbf{X} = \mathbf{AS}$ and compute its correlation matrix

$$\hat{\mathbf{R}} = \frac{1}{N}\mathbf{X}\mathbf{X}^{\mathrm{H}} = \mathbf{A}(\frac{1}{N}\mathbf{S}\mathbf{S}^{\mathrm{H}})\mathbf{A}^{\mathrm{H}} = \mathbf{A}\hat{\mathbf{R}}_s\mathbf{A}^{\mathrm{H}}$$

- Eigenvalue decomposition: $\hat{\mathbf{R}} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}}$

- *Rank property:*

  If the number of sources d is smaller than the number of antennas M

  ➠ $\boldsymbol{\Lambda}$ has d eigenvalues unequal to 0 and $M-d$ equal to zero.

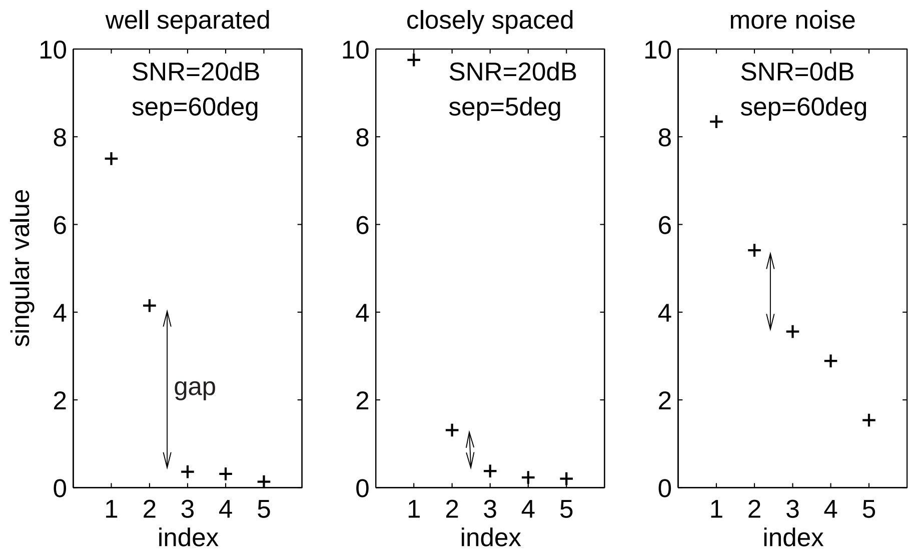
- Add i.i.d. noise: $\mathbf{X} = \mathbf{AS} + \mathbf{E}$.

$$\begin{aligned}\hat{\mathbf{R}} = \frac{1}{N}\mathbf{X}\mathbf{X}^{\mathrm{H}} &\simeq \mathbf{A}\hat{\mathbf{R}}_s\mathbf{A}^{\mathrm{H}} + \hat{\mathbf{R}}_e \\ &\simeq \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{H}} + \sigma^2\mathbf{I} \\ &= \mathbf{U}(\boldsymbol{\Lambda} + \sigma^2\mathbf{I})\mathbf{U}^{\mathrm{H}}\end{aligned}$$

All eigenvalues are raised by σ^2, but the eigenvectors stay the same.

Eigenvalue decomposition

SVD of a data matrix



$$\mathbf{X} = \mathbf{AS} + \mathbf{E}, \qquad \mathbf{A} = [\mathbf{a}(\theta_1) \quad \mathbf{a}(\theta_2)]$$

Singular values of **X** for $d = 2$ sources, $M = 5$ antennas, $N = 10$ samples.

(a) Well separated case: large gap between signal and noise singular values,

(b) signals from close directions results in a small signal singular value,

(c) increased noise level increases noise singular values.