ECE240 Introduction to Linear Dynamical Systems
Lecture 4: Matrix Factorization

Dr. Yu Zhang
ECE Department, UC Santa Cruz

Fall 2025

Outline

Column-Row (CR) Decomposition: $\mathbf{A} = \mathbf{CR}$

Given the CR decomposition:

$$\mathbf{A} = \mathbf{CR}, \quad \mathbf{C} \in \mathbb{R}^{m \times r},\ \mathbf{R} \in \mathbb{R}^{r \times n},\ \operatorname{rank}(\mathbf{A}) = r.$$

where $\mathbf{C} =$ independent columns of $\mathbf{A}$, $\mathbf{R} =$ coefficients to rebuild all columns.

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\mathbf{CR} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \end{bmatrix} = \mathbf{A}$$

$\mathrm{Range}(\mathbf{A}) = \mathrm{Range}(\mathbf{C})$ and $\mathrm{Row}(\mathbf{A}) = \mathrm{Row}(\mathbf{R})$

1. Show that $\mathrm{Range}(\mathbf{A}) = \mathrm{Range}(\mathbf{C})$

$\Rightarrow$ For any $\mathbf{y} \in \mathrm{Range}(\mathbf{A})$, $\mathbf{y} = \mathbf{A}\mathbf{x} = (\mathbf{C}\mathbf{R})\mathbf{x} = \mathbf{C}(\mathbf{R}\mathbf{x})$. Let $\mathbf{z} = \mathbf{R}\mathbf{x}$, then $\mathbf{y} = \mathbf{C}\mathbf{z} \in \mathrm{Range}(\mathbf{C})$. So $\mathrm{Range}(\mathbf{A}) \subseteq \mathrm{Range}(\mathbf{C})$.

$\Leftarrow$ Each column of $\mathbf{C}$ is one of the columns of $\mathbf{A}$, so any linear combination of columns of $\mathbf{C}$ lies in $\mathrm{Range}(\mathbf{A})$. Thus, $\mathrm{Range}(\mathbf{C}) \subseteq \mathrm{Range}(\mathbf{A})$.

$$\boxed{\mathrm{Range}(\mathbf{A}) = \mathrm{Range}(\mathbf{C})}$$

2. Show that $\mathrm{Row}(\mathbf{A}) = \mathrm{Row}(\mathbf{R})$

$$\mathbf{A}^\top = \mathbf{R}^\top\mathbf{C}^\top \quad \Rightarrow \quad \mathrm{Range}(\mathbf{A}^\top) = \mathrm{Range}(\mathbf{R}^\top) \quad \Rightarrow \quad \boxed{\mathrm{Row}(\mathbf{A}) = \mathrm{Row}(\mathbf{R})}$$

CR Decomposition for $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$

Consider $\mathbf{A} = \mathbf{C}\mathbf{R}$, where $\mathbf{C} \in \mathbb{R}^{n\times r}$ and $\mathbf{R} \in \mathbb{R}^{r\times n}$ with $\mathrm{rank}(\mathbf{A}) = r$.

System dynamics:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} = \mathbf{C}\mathbf{R}\mathbf{x}$$

Define the reduced-order state: $\mathbf{y} := \mathbf{R}\mathbf{x} \in \mathbb{R}^r$, Then

$$\dot{\mathbf{x}} = \mathbf{C}\mathbf{y}, \qquad \dot{\mathbf{y}} = \mathbf{R}\dot{\mathbf{x}} = \mathbf{R}\mathbf{C}\mathbf{y}$$

Reduced-order system:

$$\boxed{\dot{\mathbf{y}} = (\mathbf{R}\mathbf{C})\,\mathbf{y}}$$

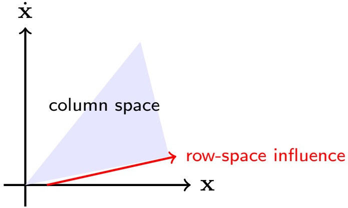
Interpretation:

- $\mathbf{C}$: basis for $\mathrm{Range}(\mathbf{A})$ — subspace where $\dot{\mathbf{x}}$ evolves.
- $\mathbf{R}$: projects $\mathbf{x}$ into the active coordinates $\mathbf{y}$
- $\mathbf{R}\mathbf{C}$: reduced-order dynamic matrix of size $r \times r$.

Dynamics View: Column Space and Row Space Roles

In the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} = \mathbf{C}\mathbf{R}\mathbf{x}$:

- Column space of $\mathbf{A}$ $(=\mathrm{Range}(\mathbf{C}))$:
  set of all possible velocity directions $\dot{\mathbf{x}}$ the system can produce.
- Row space of $\mathbf{A}$ $(=\mathrm{Row}(\mathbf{R}))$:
  set of all linear combinations of states that actually affect $\dot{\mathbf{x}}$.
- Dimensions of these spaces $(= \mathrm{rank}(\mathbf{A}))$ determine how many independent modes the system possesses.



Column space $\rightarrow$ possible motions, *Row space $\rightarrow$ active state combinations.*

LU Decomposition: $\mathbf{A} = \mathbf{LU}$

- $\mathbf{L}$: lower triangular matrix
- $\mathbf{U}$: upper triangular matrix
- Comes from Gaussian elimination
- Useful for solving $\mathbf{Ax} = \mathbf{b}$ in two steps:

$$\mathbf{Lc} = \mathbf{b}, \quad \mathbf{Ux} = \mathbf{c}$$

Why useful:

- Efficient solving of $\mathbf{Ax} = \mathbf{b}$ for equilibrium or implicit time-stepping.
- Simplifies computation of $e^{\mathbf{A}t}$ or numerical integration.
- Reveals hierarchical (triangular) dependency among states.

$\mathbf{L}$ *and* $\mathbf{U}$ *expose the causal chain of state dependencies and enable efficient computation.*

QR Decomposition: $\mathbf{A} = \mathbf{QR}$

- $\mathbf{Q}$: orthogonal matrix ($\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$); $\mathbf{R}$: upper triangular
- Columns of $\mathbf{A} \to$ orthogonal basis in $\mathbf{Q}$
- Built via Gram–Schmidt orthogonalization

Interpretation:

- $\mathbf{Q} =$ orthogonal (rotation / reflection) transformation $\to$ preserves lengths and angles.
- $\mathbf{R} =$ upper triangular scaling/shear $\to$ describes hierarchical coupling among states.

In dynamics: $\dot{\mathbf{x}} = \mathbf{Ax} = \mathbf{Q}(\mathbf{Rx})$

- $\mathbf{R}$ defines the intrinsic interaction between state components.
- $\mathbf{Q}$ rotates these effects in the state-space without changing energy.

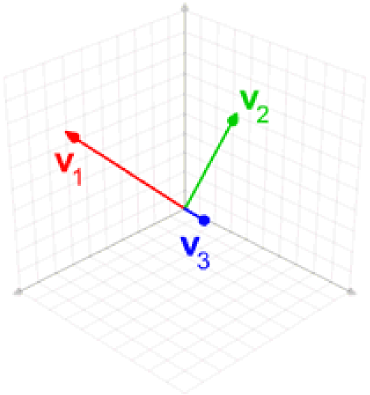
$\mathbf{Q}$: *orthogonal change of basis (energy-preserving),* $\mathbf{R}$: *directional scaling and coupling.*

Gram-Schmidt (GS) Orthogonalization

Given k nonzero linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, the GS process constructs orthogonal vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$, and orthonormal vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_k$ as follows:

$$
\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \\
\mathbf{u}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2), & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \\
\mathbf{u}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_3), & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}, \\
&\vdots & &\vdots \\
\mathbf{u}_k &= \mathbf{v}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}(\mathbf{v}_k), & \mathbf{e}_k &= \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.
\end{aligned}
$$

Gram–Schmidt Animation



Why QR Matters

1. **Numerical Stability**
   - Solving $\mathbf{A}\mathbf{x} = \mathbf{b}$: QR avoids numerical instability from ill-conditioning.
   - Used in least-squares and Kalman filtering.
2. **In dynamical systems $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$:**
   - Orthogonal basis $\mathbf{Q}$ preserves state energy and norm.
   - Enables stable time integration and modal analysis.
   - Continuous QR methods track stability and Lyapunov exponents.
3. **Physical view:**
   - $\mathbf{Q} =$ energy-preserving coordinate change.
   - $\mathbf{R} =$ internal scaling or feedforward structure.

QR links algebraic stability (orthogonality) to dynamic stability (bounded growth).

Rank-Revealing (Pivoted) QR Decomposition

Definition:

$$\mathbf{A}\boldsymbol{\Pi} = \mathbf{Q}\begin{bmatrix}\mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{R}_{22}\end{bmatrix}, \quad \mathbf{A} \in \mathbb{R}^{m\times n}.$$

- $\boldsymbol{\Pi}$: column permutation (pivoting).
- $\mathbf{Q} \in \mathbb{R}^{m\times m}$: orthogonal ($\mathbf{Q}^\top\mathbf{Q} = \mathbf{I}$).
- $\mathbf{R}_{11} \in \mathbb{R}^{r\times r}$: upper-triangular, well-conditioned.
- $\|\mathbf{R}_{22}\|$ small $\Rightarrow$ reveals numerical rank r.

Key idea: Detects linearly independent columns of $\mathbf{A}$ stably, even when $\mathbf{A}$ is nearly rank deficient.

$$\boxed{\text{RRQR} \approx \text{fast, stable rank test (like SVD but cheaper).}}$$

Rank-truncated RRQR

If the numerical rank is r, we can truncate the small trailing block $\mathbf{R}_{22} \approx \mathbf{0}$:

$$\mathbf{A} \approx [\mathbf{Q}_1 \ \mathbf{Q}_2] \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \boldsymbol{\Pi}^\top,$$

then

- Columns of $\mathbf{Q}_1$ span $\mathrm{range}(\mathbf{A})$
- Columns of $\mathbf{Q}_2$ span $\mathrm{null}(\mathbf{A}^\top)$.
- Recall that $\mathbb{R}^m = \mathrm{range}(\mathbf{A}) \oplus \mathrm{null}(\mathbf{A}^\top)$.

Modified Gram–Schmidt (MGS)

- Re-orthogonalizes step-by-step to reduce roundoff sensitivity versus classic GS.
- Same algebraic result ($\mathbf{A} = \mathbf{QR}$), better numerical behavior.
- Often combined with column pivoting for RRQR.

Projection View

$\mathbf{Q}$ spans $\mathrm{range}(\mathbf{A})$ and $\mathrm{Proj}_{\mathbf{Q}} = \mathbf{Q}\mathbf{Q}^\top$ is the orthogonal projector onto $\mathrm{range}(\mathbf{A})$.

Least Squares via QR

Solve $\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$ (full column rank):

$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 = \|\mathbf{Q}\mathbf{R}\mathbf{x} - \mathbf{b}\|_2 = \|\mathbf{Q}^\top \mathbf{b} - \mathbf{R}\mathbf{x}\|_2,$$

since $\mathbf{Q}$ is orthogonal. Partition $\mathbf{Q} = [\mathbf{Q}_1 \ \mathbf{Q}_2]$, then the solution satisfies

$$\mathbf{R}\mathbf{x} = \mathbf{Q}_1{}^\top \mathbf{b}, \quad \text{solve by back substitution.}$$

Eigenvalue Decomposition (EVD) for Square Matrices: $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$

Definition (Eigenpair and Eigenvalue Condition)

A scalar–vector pair $(\lambda, \mathbf{v})$ is an **eigenpair** of $\mathbf{A} \in \mathbb{F}^{n\times n}$ if

$$\boxed{\mathbf{A}\mathbf{v} = \lambda\mathbf{v}}, \qquad \mathbf{v} \neq \mathbf{0}.$$

Equivalently,

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}.$$

A nontrivial solution $\mathbf{v} \neq \mathbf{0}$ exists if and only if $\boxed{\det(\mathbf{A} - \lambda\mathbf{I}) = 0}$.

Geometric Interpretation:

- $\mathbf{v}$ represents a direction in which $\mathbf{A}$ acts as a pure scaling by λ.
- If $\lambda > 1$: stretching; if $0 < \lambda < 1$: shrinking; if $\lambda < 0$: reflection.
- The set of all such vectors $\mathbf{v}$ forms the **eigenspace** associated with λ.

Characteristic Polynomial of a Matrix

Definition (Characteristic Polynomial)

For a square matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$, the **characteristic polynomial** is a degree-n polynomial in λ, defined as

$$\boxed{p_{\mathbf{A}}(\lambda) = \det(\lambda \mathbf{I} - \mathbf{A})}.$$

Key Properties:

- The **eigenvalues** of $\mathbf{A}$ are the roots of $p_{\mathbf{A}}(\lambda)$: $\boxed{p_{\mathbf{A}}(\lambda_i) = 0 \quad \Longleftrightarrow \quad \det(\lambda_i \mathbf{I} - \mathbf{A}) = 0}$.
- Coefficients of $p_{\mathbf{A}}(\lambda)$ relate to invariants of $\mathbf{A}$:

$$p_{\mathbf{A}}(\lambda) = \lambda^n - \operatorname{tr}(\mathbf{A})\lambda^{n-1} + \cdots + (-1)^n \det(\mathbf{A}).$$

Geometric Interpretation:

- The roots λ_i capture the **scaling factors** of the linear map $\mathbf{A}$.
- The polynomial encodes all eigenvalue information that is the “DNA” of the matrix.

Examples

Example 1:

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \Rightarrow p_{\mathbf{A}}(\lambda) = \det\left(\begin{bmatrix} \lambda - 4 & -2 \\ -1 & \lambda - 3 \end{bmatrix}\right) = \lambda^2 - 7\lambda + 10 \quad \Rightarrow \quad \lambda_1 = 5,\ \lambda_2 = 2$$

Example 2:

$$\mathbf{A} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \implies p_{\mathbf{A}}(\lambda) = \det\left(\begin{bmatrix} \lambda & 1 & 0 \\ -1 & \lambda & 0 \\ 0 & 0 & \lambda - 2 \end{bmatrix}\right) = (\lambda^2 + 1)(\lambda - 2).$$

$$\implies \lambda_1 = 2, \qquad \lambda_{2,3} = \pm j.$$

Diagonalizable Matrices

- A matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is **diagonalizable** if it has n linearly independent eigenvectors, i.e.,

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}, \quad \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

- For a symmetric matrix:

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^\top$$

where $\mathbf{U}$ is orthonormal and $\boldsymbol{\Lambda}$ is diagonal and real.

Matrices that are Always Diagonalizable

Matrix Type	Field	Reason / Property
Symmetric $(\mathbf{A} = \mathbf{A}^\top)$	$\mathbb{R}$	Spectral theorem: $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\top$, orthogonally diagonalizable. Eigenvalues are real.
Hermitian $(\mathbf{A} = \mathbf{A}^\mathsf{H})$	$\mathbb{C}$	Unitarily diagonalizable: $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\mathsf{H}$, eigenvalues are real.
Unitary $(\mathbf{A}^\mathsf{H}\mathbf{A} = \mathbf{A}\mathbf{A}^\mathsf{H} = \mathbf{I})$	$\mathbb{C}$	Normal matrices $\Rightarrow$ unitarily diagonalizable. Eigenvalues satisfy $\lvert\lambda_i\rvert = 1$.
Normal $(\mathbf{A}^\mathsf{H}\mathbf{A} = \mathbf{A}\mathbf{A}^\mathsf{H})$	$\mathbb{C}$	All normal matrices are unitarily diagonalizable (includes Hermitian, skew-Hermitian, unitary).
Distinct Eigenvalues	$\mathbb{F}$	n distinct eigenvalues $\Rightarrow$ n linearly independent eigenvectors.

Left and Right Eigenvectors

- Right eigenvector: $\mathbf{A}\mathbf{v}_R = \lambda \mathbf{v}_R$.
- Left eigenvector: $\mathbf{v}_L{}^\top \mathbf{A} = \lambda \mathbf{v}_L{}^\top$.
- Left and right eigenvectors are related by:

$$\mathbf{v}_{\mathbf{L},i}{}^\top \mathbf{v}_{\mathbf{R},j} = 0 \quad \text{if } i \neq j.$$

- Important in dynamical systems:
  - Right eigenvectors $\rightarrow$ mode shapes of the state.
  - Left eigenvectors $\rightarrow$ sensitivity of outputs or adjoint systems.

The Role of $(\mathbf{A} - \lambda \mathbf{I})$

Name	**Meaning / Context**
Shifted matrix	$\mathbf{A}$ shifted by $\lambda \mathbf{I}$; used in algorithms.
Characteristic operator	Acts as $(\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = 0$ for eigenvectors.
Eigenmatrix (informal)	Defines the (generalized) eigenspace.

Eigenspace

Definition

For a square matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$ and an eigenvalue λ, the **eigenspace** associated with λ is

$$\mathcal{E}_\lambda = \{\, \mathbf{v} \in \mathbb{F}^n \mid \mathbf{A}\mathbf{v} = \lambda \mathbf{v} \,\} = \mathrm{Null}(\mathbf{A} - \lambda \mathbf{I}).$$

- $\mathcal{E}_\lambda$ is the subspace spanned by all eigenvectors associated with λ plus the zero vector.
- $\mathcal{E}_\lambda$ is the collection of all directions in which the linear transformation $\mathbf{A}$ acts as a simple scaling by λ.

Example

Consider $\mathbf{A} = \mathrm{diag}(2, 3)$, we have

$$\lambda = 2: \;\; \mathcal{E}_2 = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}; \quad \lambda = 3: \;\; \mathcal{E}_3 = \mathrm{span}\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

Similarity Transformation and Coordinate Change

Definition (Similarity Transformation)

Two matrices $\mathbf{A}, \mathbf{B} \in \mathbb{F}^{n\times n}$ are said to be **similar** if there exists an invertible matrix $\mathbf{P}$ such that

$$\boxed{\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}}.$$

Key Properties: $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic polynomial. Therefore

$$\lambda(\mathbf{B}) = \lambda(\mathbf{A}), \quad \det(\mathbf{B}) = \det(\mathbf{A}), \quad \mathrm{trace}(\mathbf{B}) = \mathrm{trace}(\mathbf{A}).$$

Theorem (Eigenspace relationship under similarity transformation)

If $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, *then for each eigenvalue* λ*:* $\mathcal{E}_\lambda(\mathbf{B}) = \mathbf{P}^{-1}\mathcal{E}_\lambda(\mathbf{A})$.

Proof: If $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ and $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$, then $\mathbf{B}\mathbf{y} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{y} = \lambda\mathbf{y}$. Thus, $\mathbf{y}$ is an eigenvector of $\mathbf{B}$ for the same λ.

Similarity Transformation for Linear Dynamical Systems

For a linear system: $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, consider *change coordinates* via $\mathbf{x} = \mathbf{P}\mathbf{z}$. We get

$$\dot{\mathbf{z}} = \mathbf{P}^{-1}\dot{\mathbf{x}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{z} = \mathbf{B}\mathbf{z}.$$

Thus, the system matrix in the new coordinates is **similar** to the original:

$$\boxed{\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.}$$

If $\mathbf{P}$ contains eigenvectors of $\mathbf{A}$, then $\mathbf{B} = \mathbf{\Lambda}$ (diagonal).

Now each component $z_i(t)$ evolves **independently**:

$$\dot{z}_i = \lambda_i z_i \quad \Longrightarrow \quad z_i(t) = e^{\lambda_i t} z_i(0).$$

Interpretation:

- Similarity $=$ change of coordinates (new basis).
- The underlying dynamics and eigenvalues remain invariant.
- Choosing $\mathbf{P}$ as eigenvectors $\Rightarrow$ diagonal $\mathbf{B}$ $\Rightarrow$ decoupled modal dynamics.

Singular Value Decomposition (SVD): $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top$

Extending eigenvalue decomposition for non-square matrices, we get full SVD

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\mathbf{V}_{n\times n}^\top,$$

where

- $\mathbf{U}$ and $\mathbf{V}$ are orthonormal: $\mathbf{U}\mathbf{U}^\top = \mathbf{U}^\top\mathbf{U} = \mathbf{I}$, $\mathbf{V}\mathbf{V}^\top = \mathbf{V}^\top\mathbf{V} = \mathbf{I}$.
- $\boldsymbol{\Sigma}$ is a rectangular diagonal matrix whose non-negative diagonal entries σ_i represent the singular values of $\mathbf{A}$.
- $\sigma_i = \sqrt{\lambda_i(\mathbf{A}\mathbf{A}^\top)}$. If $\mathbf{A}$ is symmetric, then $\sigma_i(\mathbf{A}) = |\lambda_i(\mathbf{A})|$.
- *Rank of a matrix*: $r =$ number of non-zero singular values.
- Columns of $\mathbf{U}$ are *left singular vectors*, which are the eigenvectors of $\mathbf{A}\mathbf{A}^\top$.
- Columns of $\mathbf{V}$ are *right singular vectors*, which are the eigenvectors of $\mathbf{A}^\top\mathbf{A}$.

Compact SVD

Expresses $\mathbf{A}$ as sum of rank-1 components:

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times r}\boldsymbol{\Sigma}_{r\times r}\mathbf{V}_{r\times n}^\top = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$$

where

- $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_r]$ and $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_r]$ are column-orthonormal: $\mathbf{U}^\top\mathbf{U} = \mathbf{V}^\top\mathbf{V} = \mathbf{I}_r$
- $\boldsymbol{\Sigma}$ is a square diagonal matrix with all positive singular values of $\mathbf{A}$ on the diagonal.

SVD and Four Subspaces (1/3)

Consider the SVD of a given matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with rank r,

$$\mathbf{A}_{m\times n} = \mathbf{U}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\mathbf{V}_{n\times n}^\top = [\mathbf{U}_L, \mathbf{U}_R]\begin{bmatrix}\mathbf{D}_{r\times r} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}[\mathbf{V}_L, \mathbf{V}_R]^\top$$

$$= [\mathbf{u}_1, \ldots, \mathbf{u}_r, \mathbf{u}_{r+1}, \ldots, \mathbf{u}_m]\begin{bmatrix}\mathbf{D}_{r\times r} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{v}_1^\top\\ \vdots\\ \mathbf{v}_r^\top\\ \mathbf{v}_{r+1}^\top\\ \vdots\\ \mathbf{v}_n^\top\end{bmatrix}$$

where $\mathbf{D}_{r\times r} = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ with ordered singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

SVD and Four Subspaces (2/3)

The full SVD provides basis vectors of 4 subspaces associated with $\mathbf{A}$:

- Column space: $\mathcal{R}(\mathbf{A}) = \mathcal{R}(\mathbf{U}_L) = \text{span}\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$
- Row space: $\mathcal{R}(\mathbf{A}^\top) = \mathcal{R}(\mathbf{V}_L) = \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$
- Left null space: $\mathcal{N}(\mathbf{A}^\top) = \mathcal{R}(\mathbf{U}_R) = \text{span}\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\}$
- Null space: $\mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{V}_R) = \text{span}\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$

SVD and Four Subspaces (3/3)

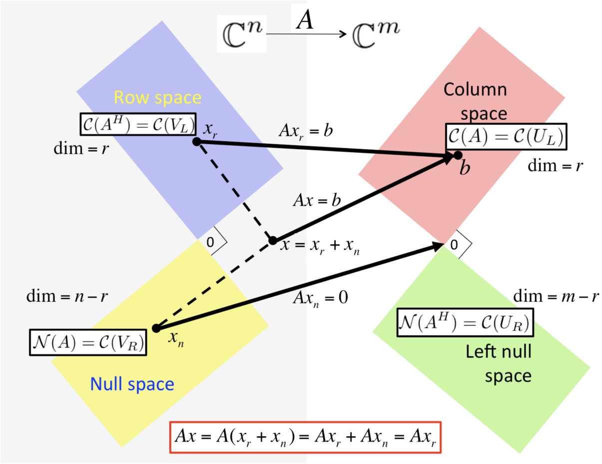


Figure: Source link

Matrix Factorization Matlab Commands

Factorization	Mathematical Form	MATLAB Command
CR	$\mathbf{A} = \mathbf{CR}$	`no built-in command`[1]
LU	$\mathbf{PA} = \mathbf{LU}$	`[L,U,P]=lu(A)`
QR	$\mathbf{A} = \mathbf{QR}$	`[Q,R]=qr(A)`
EVD	$\mathbf{A} = \mathbf{V\Lambda V}^{-1}$	`[V,D]=eig(A)`
SVD	$\mathbf{A} = \mathbf{U\Sigma V}^{\top}$	`[U,S,V]=svd(A)`
Jordan	$\mathbf{A} = \mathbf{PJP}^{-1}$	`[V,J]=jordan(A)`

Table: Summary of matrix factorizations and corresponding MATLAB commands; see more factorizations in this reference link

[1] by pivoted QR: [Q,R,P]= qr(A,0); r = rank(A); C = A(:,P(1:r)); R = C\A

Positive Definite Matrices

- If all eigenvalues are positive, symmetric $\mathbf{A}$ is *positive definite* $\mathbf{A} \succ \mathbf{0}$.
- If all eigenvalues are non-negative, symmetric $\mathbf{A}$ is *positive semi-definite* $\mathbf{A} \succeq \mathbf{0}$.

- $\mathbf{A} \succeq \mathbf{0}$ iff all principal minors of $\mathbf{A}$ are nonnegative.
- $\mathbf{A} \succeq \mathbf{0}$ iff there exists $\mathbf{B}$ such that $\mathbf{A} = \mathbf{B}^\top \mathbf{B}$.
- $\mathbf{A} \succeq \mathbf{0}$ iff $\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0$ for all $\mathbf{x}$.
- *Square root*: $\mathbf{A}^{1/2} = \mathbf{U}\sqrt{\mathbf{\Lambda}}\mathbf{U}^\top$

Lyapunov Stability: Core Idea

- Lyapunov's method studies stability using an **energy-like function**:

$$V(\mathbf{x}) = \mathbf{x}^\top \mathbf{P}\mathbf{x},$$

  where $\mathbf{P}$ is a **symmetric positive definite** matrix $(\mathbf{P} \succ 0)$.

- $V(\mathbf{x})$ acts like a potential or energy function:

$$V(\mathbf{x}) > 0 \text{ for } \mathbf{x} \neq \mathbf{0}, \quad V(\mathbf{0}) = 0.$$

- The sign of its time derivative $\dot{V}$ determines whether energy decreases over time.

Continuous-Time Linear System

Consider a linear system:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}.$$

The time derivative of $V(\mathbf{x}) = \mathbf{x}^\top \mathbf{P}\mathbf{x}$ is

$$\dot{V} = \mathbf{x}^\top (\mathbf{A}^\top \mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{x}.$$

If there exists $\mathbf{P} \succ 0$ such that

$$\mathbf{A}^\top \mathbf{P} + \mathbf{P}\mathbf{A} \prec 0,$$

then $\dot{V} < 0$ for all $\mathbf{x} \neq \mathbf{0}$, implying **asymptotic stability**. This condition is equivalent to the **Lyapunov equation**:

$$\mathbf{A}^\top \mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q}, \quad \mathbf{Q} \succ 0.$$

Relation to Definiteness

- **Positive definite $\mathbf{P}$** ensures $V(\mathbf{x}) > 0$ (energy is always positive).
- **Negative definite $\dot{V}$** ensures $V(\mathbf{x})$ decreases over time.
- Together, these guarantee that trajectories decay to the equilibrium $\mathbf{x} = \mathbf{0}$.

$$\mathbf{P} \succ 0 \quad \text{and} \quad \mathbf{A}^\top\mathbf{P} + \mathbf{PA} \prec 0 \quad \Rightarrow \quad \text{Asymptotic stability.}$$

Interpretation: Matrix definiteness provides an algebraic test for energy dissipation and stability in linear dynamical systems.

Schur Complement[2]

- For invertible $\mathbf{A}$ and matrix $\mathbf{X} \in \mathbb{S}^N$: $\mathbf{X} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix}$
- Define **Schur complement** as $\mathbf{S} := \mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}$
- Property 1: $\det(\mathbf{X}) = \det(\mathbf{A}) \det(\mathbf{S})$
- Property 2: $\mathbf{X} \succ \mathbf{0}$ iff $\mathbf{A} \succ \mathbf{0}$ and $\mathbf{S} \succ \mathbf{0}$
- Property 3: For $\mathbf{A} \succ \mathbf{0}$ it holds, $\mathbf{X} \succeq \mathbf{0}$ iff $\mathbf{S} \succeq \mathbf{0}$
- For linear dynamical systems: Schur complement is useful in *state-space reduction*, *Lyapunov and Riccati equations*, *block matrix inversion*, *and stability/definiteness tests*.

[2] see page 672 of the book *Convex Optimization* by Boyd-Vandenberghe