Lec 8: Singular Value Decomposition

15-369/669/769: Numerical Computing

Instructor: Minchen Li

Table of Content

Table of Content

Deriving the SVD

Linear Map Viewpoint

- Consider $A \in \mathbb{R}^{m \times n}$ as a map
$$\mathbf{v} \mapsto A\mathbf{v} : \mathbb{R}^n \to \mathbb{R}^m.$$
- We study how lengths change:
$$R(\mathbf{v}) = \frac{\|A\mathbf{v}\|_2}{\|\mathbf{v}\|_2}, \qquad \mathbf{v} \neq 0.$$
- Scale invariance:
$$R(\alpha\mathbf{v}) = \frac{\|A(\alpha\mathbf{v})\|_2}{\|\alpha\mathbf{v}\|_2} = \frac{|\alpha|\ \|A\mathbf{v}\|_2}{|\alpha|\ \|\mathbf{v}\|_2} = R(\mathbf{v}).$$
- Therefore, we may restrict to unit vectors: $\|\mathbf{v}\|_2 = 1$.

Deriving the SVD

Rayleigh-Quotient Problem

- Since $R(\mathbf{v}) \geq 0$,

$$[R(\mathbf{v})]^2 = \|A\mathbf{v}\|_2^2 = (A\mathbf{v})^\top (A\mathbf{v}).$$

- Expand:

$$(A\mathbf{v})^\top (A\mathbf{v}) = \mathbf{v}^\top A^\top A\mathbf{v}.$$

- Thus, finding extremal values of $R(\mathbf{v})$ is equivalent to finding extremal values of the quadratic form

$$\mathbf{v}^\top A^\top A\mathbf{v} \quad \text{subject to } \|\mathbf{v}\|_2 = 1.$$

- Critical points are eigenvectors:

$$A^\top A\,\mathbf{v}_i = \lambda_i \mathbf{v}_i, \qquad \lambda_i \geq 0.$$

Deriving the SVD

Relating $A^T A$ and AA^T

- Define

$$\mathbf{u}_i^* := A\mathbf{v}_i.$$

- If $\lambda_i \neq 0$, then

$$AA^\top \mathbf{u}_i^* = A(A^\top A\mathbf{v}_i) = A(\lambda_i \mathbf{v}_i) = \lambda_i(A\mathbf{v}_i) = \lambda_i \mathbf{u}_i^*.$$

- So $\mathbf{u}_i^*$ is an eigenvector of $AA^\top$ with eigenvalue λ_i.

- Norm:

$$\|\mathbf{u}_i^*\|_2^2 = \|A\mathbf{v}_i\|_2^2 = \mathbf{v}_i^\top (A^\top A)\mathbf{v}_i = \lambda_i \|\mathbf{v}_i\|_2^2 = \lambda_i.$$

- Therefore, if $\lambda_i > 0$, then $\|\mathbf{u}_i^*\|_2 = \sqrt{\lambda_i}$.

Deriving the SVD
Normalized Eigenvector Pairs

- Normalize $\mathbf{u}_i := \mathbf{u}_i^* / \|\mathbf{u}_i^*\|$ when $\lambda_i > 0$.
- Then
$$A\mathbf{v}_i = \sqrt{\lambda_i}\,\mathbf{u}_i, \qquad A^\top \mathbf{u}_i = \sqrt{\lambda_i}\,\mathbf{v}_i.$$
- This shows a one-to-one pairing between:
  - eigenvectors of $A^\top A$ ($\mathbf{v}_i$),
  - eigenvectors of $AA^\top$ ($\mathbf{u}_i$),
  - with the same eigenvalue λ_i.
  - Also, $\mathbf{u}_i^T A\mathbf{v}_i = (\frac{1}{\sqrt{\lambda_i}}\mathbf{v}_i^T A^T)A\mathbf{v}_i = \frac{1}{\sqrt{\lambda_i}}\mathbf{v}_i^T \lambda_i \mathbf{v}_i = \sqrt{\lambda_i}$

Deriving the SVD

Core Representation

- Let $k = \#\{i : \lambda_i > 0\}$.
- Collect eigenvectors:

$$\bar{V} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k] \in \mathbb{R}^{n \times k}, \quad \bar{U} = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_k] \in \mathbb{R}^{m \times k}.$$

- Define diagonal matrix

$$\bar{\Sigma} = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_k}).$$

- Then

$$\bar{U}^\top A \bar{V} = \bar{\Sigma}.$$

Deriving the SVD

Extending to Full SVD

- Extend $\bar{U}$, $\bar{V}$ to orthogonal bases:

$$U = [\bar{U}\ U_0] \in \mathbb{R}^{m \times m}, \qquad V = [\bar{V}\ V_0] \in \mathbb{R}^{n \times n}.$$

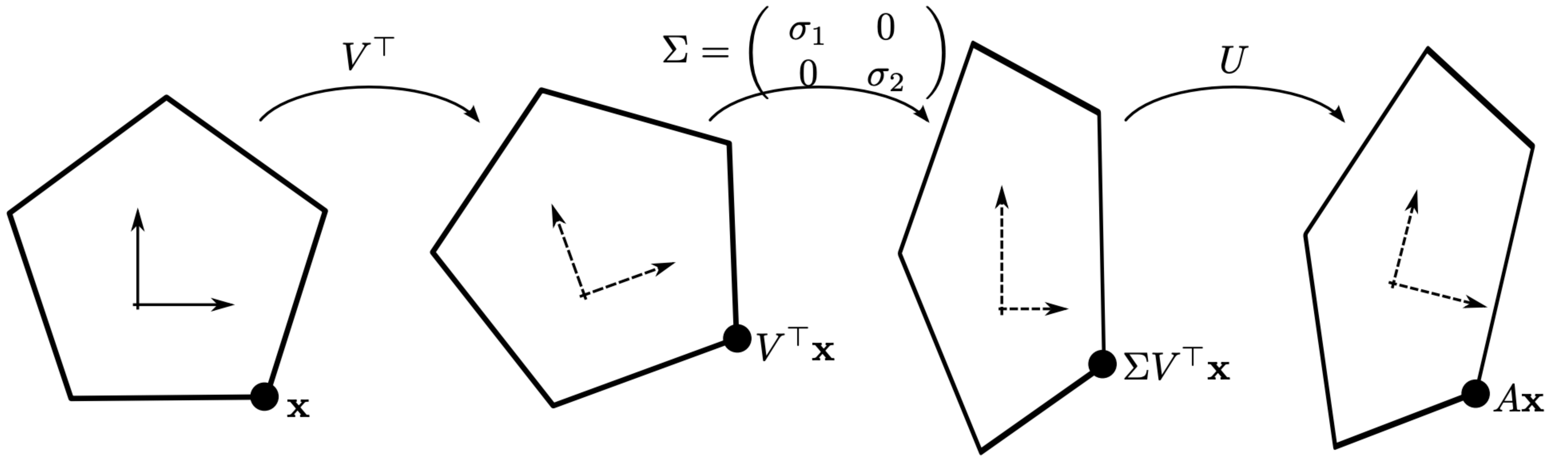
- Define $\Sigma \in \mathbb{R}^{m \times n}$ by

$$\Sigma_{ij} = \begin{cases} \sqrt{\lambda_i}, & i = j \le k, \\ 0, & \text{otherwise.} \end{cases}$$

- Orthogonality gives

$$U^\top A V = \Sigma, \qquad \Longleftrightarrow \qquad A = U \Sigma V^\top.$$

Deriving the SVD
Interpretation and Nomenclature



- Columns of U: *left singular vectors*.
- Columns of V: *right singular vectors*.
- Diagonal entries of Σ: *singular values* $\sigma_i = \sqrt{\lambda_i}$.
- Ordered convention: $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$.

- Geometric interpretation:
  - V: isometry in $\mathbb{R}^n$.
  - Σ: axis-aligned scaling by σ_i.
  - U: isometry in $\mathbb{R}^m$.

Deriving the SVD

Algebraic Derivation

- We can also algebraically derive the SVD of A from the eigendecomposition of $B = \begin{bmatrix} & A^T \\ A & \end{bmatrix}$

> **Proposition 7.1.** Take $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{R}^{m+n}$ to be an eigenvector of B defined above with eigenvalue λ, where $\mathbf{x}_1 \in \mathbb{R}^n$ and $\mathbf{x}_2 \in \mathbb{R}^m$. Then, $\mathbf{x}' := (\mathbf{x}_1, -\mathbf{x}_2)$ is an eigenvector of B with eigenvalue $-\lambda$.

- Take $\Sigma \in \mathbb{R}^{k\times k}$ to be a diagonal matrix containing the positive eigenvalues of B, and take the columns of $X \in \mathbb{R}^{(m+n)\times k}$ to be the corresponding eigenvectors: $X = \begin{bmatrix} V \\ U \end{bmatrix}$, where $V \in \mathbb{R}^{n\times k}$ and $U \in \mathbb{R}^{m\times k}$.

- Then $-\Sigma$ and $X' = \begin{bmatrix} V \\ -U \end{bmatrix}$ are the negative eigenvalues and the corresponding eigenvectors.

- For the remaining zero eigenvalues, similarly denote their eigenvectors $N = \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}$

Deriving the SVD

Algebraic Derivation and SVD Computation

- With the previous setup, we can derive

$$\begin{pmatrix} 0 & 2A^\top \\ 2A & 0 \end{pmatrix} = \begin{pmatrix} V & V & N_1 \\ U & -U & N_2 \end{pmatrix} \begin{pmatrix} \Sigma & 0 & 0 \\ 0 & -\Sigma & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} V^\top & U^\top \\ V^\top & -U^\top \\ N_1^\top & N_2^\top \end{pmatrix} = \begin{pmatrix} 0 & 2V\Sigma U^\top \\ 2U\Sigma V^\top & 0 \end{pmatrix}.$$

$$V^\top V = U^\top U = I_{k\times k}$$

- This derivation also provides us a way to compute the SVD of A via eigendecomposition on $B = \begin{bmatrix} & A^T \\ A & \end{bmatrix}$ without explicitly forming $A^T A$.

- Practical methods for computing SVD first convert A to a bidiagonal matrix, and then apply iterative methods, e.g. Jacobi, to compute the SVD.

Table of Content

Solving Linear Systems and the Pseudoinverse

Solving $A\mathbf{x} = \mathbf{b}$ using SVD (Square & Invertible)

- Factor $A = U\Sigma V^\top$ with $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ orthogonal, $\Sigma = \mathsf{diag}(\sigma_1, \dots, \sigma_n)$ with $\sigma_i > 0$.
- Multiply $U^\top$ on the left: $U^\top A\mathbf{x} = U^\top\mathbf{b} \Rightarrow \Sigma V^\top\mathbf{x} = \mathbf{c}$ with $\mathbf{c} := U^\top\mathbf{b}$.
- Multiply V on the right: $V^\top\mathbf{x} = \Sigma^{-1}\mathbf{c} \Rightarrow \mathbf{x} = V\Sigma^{-1}U^\top\mathbf{b}$.
- Here $\Sigma^{-1} = \mathsf{diag}(1/\sigma_1, \dots, 1/\sigma_n)$.

Solving Linear Systems and the Pseudoinverse

Least-Squares When A is Not Square/Invertible

- **Goal:** solve $A\mathbf{x} \approx \mathbf{b}$ with minimal residual $\|A\mathbf{x} - \mathbf{b}\|^2$.
- **Normal equations:** at any least-squares minimizer, $A^\top(A\mathbf{x} - \mathbf{b}) = \mathbf{0} \Rightarrow A^\top A\mathbf{x} = A^\top\mathbf{b}$.
- When $A^\top A$ is invertible (full column rank), the solution is unique: $\mathbf{x} = (A^\top A)^{-1}A^\top\mathbf{b}$.
- Underdetermined or rank-deficient: many $\mathbf{x}$ satisfy $A^\top A\mathbf{x} = A^\top\mathbf{b}$. Prefer the *minimum-norm* solution:

$$\min_{\mathbf{x}} \|\mathbf{x}\|^2 \quad \text{s.t.} \quad A^\top A\mathbf{x} = A^\top\mathbf{b}.$$

Solving Linear Systems and the Pseudoinverse

Rewriting the Constraint using $A = U\Sigma V^T$

- Compute $A^\top A = (U\Sigma V^\top)^\top (U\Sigma V^\top) = V\Sigma^\top \Sigma V^\top = V\Sigma^2 V^\top$ (since $U^\top U = I$).
- Constraint $A^\top A\mathbf{x} = A^\top \mathbf{b}$ becomes

$$V\Sigma^2 V^\top \mathbf{x} = V\Sigma U^\top \mathbf{b}.$$

- **Change variables:** $\mathbf{y} := V^\top \mathbf{x}$ and $\mathbf{d} := U^\top \mathbf{b}$ (orthogonal transforms preserve ℓ_2-norm).
- Then the constrained minimum-norm problem is

$$\min_{\mathbf{y}} \ \|\mathbf{y}\|^2 \quad \text{s.t.} \quad \Sigma^2 \mathbf{y} = \Sigma \mathbf{d}.$$

Solving Linear Systems and the Pseudoinverse

Diagonal Decoupling in Σ-Coordinates

- $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_\ell)$ with $\ell = \min\{m, n\}$; write constraints componentwise:

$$\sigma_i^2 \, y_i = \sigma_i \, d_i \quad \text{for } i = 1, \ldots, \ell.$$

- For $\sigma_i > 0$: $y_i = d_i / \sigma_i$.
- For $\sigma_i = 0$: the constraint imposes no condition; to minimize $\|\mathbf{y}\|_2$, choose $y_i = 0$.
- Therefore the minimizer is $\mathbf{y} = \Sigma^+ \mathbf{d}$, where the *diagonal* Σ^+ is

$$(\Sigma^+)_{ii} = \begin{cases} 1/\sigma_i, & \sigma_i > 0, \\ 0, & \sigma_i = 0. \end{cases}$$

- Undo the variables: $\mathbf{x} = V\mathbf{y} = V\Sigma^+ U^\top \mathbf{b}$.

Solving Linear Systems and the Pseudoinverse

Moore–Penrose Pseudoinverse

- **Definition:** For $A = U\Sigma V^\top$, the pseudoinverse is

$$A^+ := V\Sigma^+ U^\top \in \mathbb{R}^{n\times m}.$$

- $\mathbf{x}^\star = A^+\mathbf{b}$ is the *minimum-norm* vector among all $\mathbf{x}$ satisfying the normal equations.
- Special cases:
  - If A is square and invertible, $A^+ = A^{-1}$.
  - If A is overdetermined (full column rank), $A^+ = (A^\top A)^{-1}A^\top$ gives the unique least-squares solution.
  - If A is underdetermined (full row rank), $A^+ = A^\top(AA^\top)^{-1}$ gives the minimum-norm solution.

Table of Content

Low-Rank Approximations

Outer-Product Expansion and Fast Application

- Using SVD, expand A as a sum of rank-1 terms:

$$A = \sum_{i=1}^{\ell} \sigma_i \, \mathbf{u}_i \mathbf{v}_i^\top, \qquad \ell = \min\{m, n\}.$$

- Action on a vector:

$$A\mathbf{x} = \sum_{i=1}^{\ell} \sigma_i \, \mathbf{u}_i \, (\mathbf{v}_i^\top \mathbf{x}) = \sum_{i=1}^{\ell} \sigma_i \, (\mathbf{v}_i \cdot \mathbf{x}) \, \mathbf{u}_i.$$

- **Interpretation:** project $\mathbf{x}$ onto right singular directions $\mathbf{v}_i$, scale by σ_i, re-express along left singular directions $\mathbf{u}_i$.

Low-Rank Approximations

Truncation and Low-Rank Approximation

- If many σ_i are small, approximate

$$A \approx A_k := \sum_{i=1}^{k} \sigma_i \, \mathbf{u}_i \mathbf{v}_i^\top \quad (k < \ell).$$

- Then $A\mathbf{x}$ can be approximated using only the top k terms:

$$A\mathbf{x} \approx \sum_{i=1}^{k} \sigma_i \, (\mathbf{v}_i^\top \mathbf{x}) \, \mathbf{u}_i .$$

- Pseudoinverse also truncates: $A^+ = \sum_{\sigma_i \neq 0} \frac{\mathbf{v}_i \mathbf{u}_i^\top}{\sigma_i} \;\approx\; \sum_{i=1}^{k} \frac{\mathbf{v}_i \mathbf{u}_i^\top}{\sigma_i}$.

- **Practical benefit:** compute/apply A_k or A_k^+ using only leading singular triplets.

Low-Rank Approximations

Eckart–Young Theorem (Optimality of Truncation)

- Let A_k be obtained by zeroing all but the largest k singular values of A.
- **Theorem (Eckart–Young, 1936).** For both spectral norm and Frobenius norm,

$$A_k = \arg \min_{\mathrm{rank}(B) \leq k} \|A - B\|.$$

- In spectral norm, the optimal error equals the next singular value:

$$\|A - A_k\|_2 = \sigma_{k+1}.$$

Low-Rank Approximations

Eckart–Young (Sketch Proof, Spectral Norm)

- By SVD and definition of A_k: $A - A_k = \sum_{i=k+1}^{\ell} \sigma_i \, \mathbf{u}_i \mathbf{v}_i^\top, \quad \Rightarrow \quad \|A - A_k\|_2 = \sigma_{k+1}$.
- For any rank-k matrix $B_k = XY^\top$ ($X \in \mathbb{R}^{m\times k}$, $Y \in \mathbb{R}^{n\times k}$):
  - Let $V_{k+1:\ell} = [\mathbf{v}_{k+1} \ \cdots \ \mathbf{v}_\ell]$.
  - Because $\mathrm{rank}(Y^\top V_{k+1:\ell}) < \ell - k$, there exists $\mathbf{z} \neq 0$ with $Y^\top V_{k+1:\ell}\mathbf{z} = 0$; rescale to $\|\mathbf{z}\|_2 = 1$.
  - Take $\mathbf{q} := V_{k+1:\ell}\mathbf{z}$ (unit). Then $\|(A - B_k)\mathbf{q}\|_2 = \|U\Sigma V^\top \mathbf{q} - XY^\top \mathbf{q}\|_2 = \|\Sigma V^\top \mathbf{q}\|_2$ since $Y^\top \mathbf{q} = 0$.
  - But $V^\top \mathbf{q} = [\mathbf{0}; \mathbf{z}] \Rightarrow \|\Sigma V^\top \mathbf{q}\|_2^2 = \sum_{i=k+1}^{\ell} \sigma_i^2 z_i^2 \geq \sigma_{k+1}^2$.
- Hence $\|A - B_k\|_2 \geq \sigma_{k+1} = \|A - A_k\|_2$.

Table of Content

Matrix Norms

Frobenius Norm: Definition and SVD Setup

- Frobenius norm:

$$\|A\|_{\mathrm{F}}^2 := \sum_{i,j} a_{ij}^2 \;=\; \mathrm{tr}(A^\top A).$$

- SVD of $A \in \mathbb{R}^{m\times n}$: $A = U\Sigma V^\top$, where U, V orthogonal and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_\ell)$, $\ell = \min\{m, n\}$.

- Use: $U^\top U = I$, $V^\top V = I$, cyclic trace $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Matrix Norms

Frobenius Norm and Sum of Squared Singular Values

- Compute:

$$\begin{aligned}
\|A\|_{\mathrm{F}}^2 &= \mathrm{tr}(A^\top A) \\
&= \mathrm{tr}\big((U\Sigma V^\top)^\top (U\Sigma V^\top)\big) \\
&= \mathrm{tr}\big(V\Sigma^\top U^\top U\Sigma V^\top\big) \\
&= \mathrm{tr}\big(V\Sigma^\top \Sigma V^\top\big) \\
&= \mathrm{tr}\big(\Sigma^\top \Sigma V^\top V\big) \\
&= \mathrm{tr}(\Sigma^\top \Sigma) \;=\; \sum_{j=1}^{\ell} \sigma_j^2 \quad (\Sigma \text{ diagonal}).
\end{aligned}$$

- $\Rightarrow$ Squared Frobenius norm is the sum of squared singular values.

Matrix Norms

Spectral Norm (Induced 2-Norm)

- Operator 2-norm:

$$\|A\|_2 = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2 = \sqrt{\max_{\|\mathbf{x}\|_2=1} \mathbf{x}^\top A^\top A\mathbf{x}}.$$

- Rayleigh quotient: $\max_{\|\mathbf{x}\|=1} \mathbf{x}^\top (A^\top A)\mathbf{x} = \lambda_{\max}(A^\top A)$.
- Since eigenvalues of $A^\top A$ are σ_i^2,

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^\top A)} = \max_i \sigma_i = \sigma_{\max}.$$

- Similarly, $\min\limits_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|_2 = \sigma_{\min}$ (possibly 0 if A is rank-deficient).

Matrix Norms

Condition Number and Relation to Singular Values

- For invertible A, 2-norm condition number:

$$\operatorname{cond}_2(A) := \|A\|_2 \, \|A^{-1}\|_2.$$

- SVD implies singular values of A^{-1} are reciprocals: $\sigma_i(A^{-1}) = 1/\sigma_i(A)$.
- Hence

$$\operatorname{cond}_2(A) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}.$$

- Interpretation:
  - Large ratio $\Rightarrow$ $\sigma_{\min}$ small $\Rightarrow$ ill-conditioned; small relative errors in $\mathbf{b}$ may greatly amplify in $\mathbf{x}$.
  - Well-conditioned matrices have singular values of comparable magnitude.

Matrix Norms

Practical Notes on Computing Norms and $\sigma_{\min}$

- $\|A\|_{\mathrm{F}}$ is cheap from entries; also equals $\left(\sum \sigma_i^2\right)^{1/2}$.
- $\|A\|_2 = \sigma_{\max}$: compute via power iterations on $A^\top A$ or directly via partial SVD.
- $\sigma_{\min}$ can be harder:
  - Inverse iteration/shifted methods or partial SVD near smallest singular values.
  - Solving $A\mathbf{x} = \mathbf{b}$ during iterations may itself be ill-conditioned when $\sigma_{\min}$ is tiny.
  - Use bounds/inequalities, robust factorizations (QR/SVD), or regularization if necessary.
- **Takeaway:** SVD provides clean formulas $\|A\|_{\mathrm{F}}^2 = \sum \sigma_i^2$, $\|A\|_2 = \sigma_{\max}$, $\mathrm{cond}_2(A) = \sigma_{\max}/\sigma_{\min}$, and guides numerical strategy.

Table of Content

Principal Component Analysis (PCA)

Setup

- Data matrix $X \in \mathbb{R}^{n \times k}$, k samples in n-dimensional space.
- **Goal:** find a d-dimensional subspace ($d \leq \min\{k, n\}$) capturing maximum variance in the data.
- Choose $C \in \mathbb{R}^{n \times d}$ with orthonormal columns ($C^\top C = I_d$).
- Projection of X onto subspace spanned by C: $CC^\top X$.
- Optimization problem:

$$\min_{C^\top C = I_d} \|X - CC^\top X\|_{\mathrm{F}}^2.$$

Principal Component Analysis (PCA)
Simplifying the Objective

- Expand the Frobenius norm:

$$\|X - CC^\top X\|_{\mathrm{F}}^2 = \mathrm{tr}\big((X - CC^\top X)^\top (X - CC^\top X)\big).$$

- Simplify using trace properties:

$$= \mathrm{tr}(X^\top X) - 2\,\mathrm{tr}(X^\top CC^\top X) + \mathrm{tr}(X^\top CC^\top CC^\top X).$$

- Since $C^\top C = I_d$:

$$= \mathrm{tr}(X^\top X) - \mathrm{tr}(X^\top CC^\top X).$$

- Equivalent to maximizing:

$$\|C^\top X\|_{\mathrm{F}}^2.$$

Principal Component Analysis (PCA)

PCA via SVD

- Compute SVD of data matrix:
$$X = U\Sigma V^\top.$$
- Let $\tilde{C} := U^\top C$. Then
$$\|C^\top X\|_{\mathrm{F}}^2 = \|\Sigma^\top \tilde{C}\|_{\mathrm{F}}^2.$$
- Expanding:
$$\|\Sigma^\top \tilde{C}\|_{\mathrm{F}}^2 = \sum_i \sigma_i^2 \sum_j \tilde{c}_{ij}^2.$$
- Intuitively, $\tilde{c}_{ij}$ is the projection of u_i onto c_j. Since $\|u_i\| = 1$, we have $\sum_j \tilde{c}_{ij}^2 \leq 1$, and so each σ_i^2 gets weight ≤ 1.
- Maximum achieved by aligning c_j's with u_i's corresponding to the largest singular values.

Principal Component Analysis (PCA)

Solution

- **Optimal choice:** columns of C are the top d left singular vectors of X.
- **Equivalently:** PCA directions = eigenvectors of $XX^\top$ with largest eigenvalues.
- **Intuition:** projection maximizes variance of projected data.
- **In practice:** data matrix X is often centered (columns have zero mean).

Principal Component Analysis (PCA)

Eigenfaces: PCA for Face Recognition

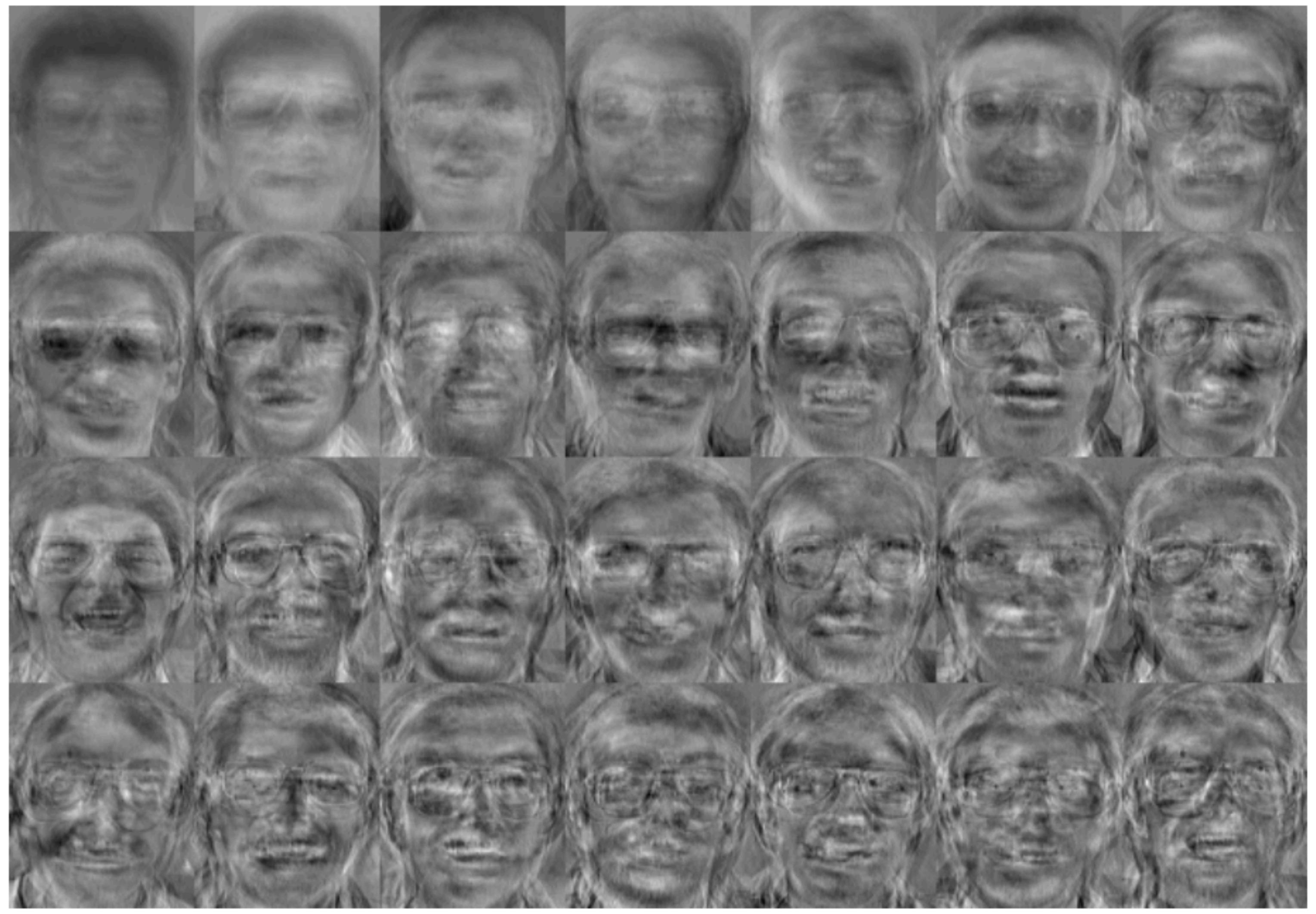
- Store training set of faces in columns of matrix $X \in \mathbb{R}^{mn \times k}$.
- Subtract mean face: center the dataset.
- Apply PCA to X, compute basis $C \in \mathbb{R}^{mn \times d}$.
- Columns of C = *eigenfaces*, capturing main modes of variation (face shape, features, lighting).
- Each face image represented as coefficient vector:

$$\mathbf{y} = C^\top \mathbf{x}.$$

Principal Component Analysis (PCA)

Eigenfaces Visualization

(a) Input faces



(b) Eigenfaces

(c) Projection

Principal Component Analysis (PCA)

Face Recognition with Eigenfaces

- For a new image $\mathbf{x}$:
  - Project: $\mathbf{y} = C^\top \mathbf{x}$.
  - Compare $\mathbf{y}$ to projections of training images $C^\top X$.
  - Closest match determines recognition result.
- Advantages:
  - Dimension reduction: $d \ll mn$.
  - Separates relevant modes (identity) from irrelevant ones (lighting, noise).

Principal Component Analysis (PCA)

Practical Notes on Eigenfaces

- Eigenfaces effective despite simplicity; e.g., training using photos of 40 subjects and then test using 40 different photos of the same subjects achieves 80% recognition accuracy.
- Real systems use enhancements:
  - Thresholds for match/no-match detection.
  - Larger basis to capture more variation.
  - Robust preprocessing (alignment, lighting normalization).
- PCA remains foundational: inspires modern subspace and feature-based recognition methods.

Table of Content