SI231 Matrix Analysis and Computations
Topic 7: Singular Value Decomposition

Ziping Zhao

Spring Term 2020–2021

School of Information Science and Technology
ShanghaiTech University, Shanghai, China

Topic 7: Singular Value Decomposition

- singular value decomposition
- matrix norms
- linear systems
- LS, pseudo-inverse, orthogonal projections
- low-rank matrix approximation
- singular value inequalities
- computation of the SVD

Main Results

- any matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ admits a singular value decomposition

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T,$$

  where $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ are orthogonal, and $\boldsymbol{\Sigma} \in \mathbb{R}^{m \times n}$ has $[\boldsymbol{\Sigma}]_{ij} = 0$ for all $i \neq j$ and $[\boldsymbol{\Sigma}]_{ii} = \sigma_i$ for all i, with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{\min\{m,n\}} \geq 0$.

- matrix 2-norm: $\|\mathbf{A}\|_2 = \sigma_1$

- let r be the number of nonzero σ_i's, partition $\mathbf{U} = [\ \mathbf{U}_1\ \mathbf{U}_2\ ]$, $\mathbf{V} = [\ \mathbf{V}_1\ \mathbf{V}_2\ ]$, and let $\tilde{\boldsymbol{\Sigma}} = \mathrm{Diag}(\sigma_1, \ldots, \sigma_r)$
  - thin SVD: $\mathbf{A} = \mathbf{U}_1\tilde{\boldsymbol{\Sigma}}\mathbf{V}_1^T$
  - pseudo-inverse: $\mathbf{A}^\dagger = \mathbf{V}_1\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{U}_1^T$
  - Linear system solution: $\mathbf{x} = \mathbf{A}^\dagger\mathbf{y} + \boldsymbol{\eta}$ for any $\boldsymbol{\eta} \in \mathcal{R}(\mathbf{V}_2), \mathbf{U}_2^T\mathbf{y} = \mathbf{0}$
  - LS solution: $\mathbf{x}_{\mathsf{LS}} = \mathbf{A}^\dagger\mathbf{y} + \boldsymbol{\eta}$ for any $\boldsymbol{\eta} \in \mathcal{R}(\mathbf{V}_2)$
  - orthogonal projection: $\mathbf{P}_{\mathbf{A}} = \mathbf{U}_1\mathbf{U}_1^T$

Main Results

- low-rank matrix approximation: given $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $k \in \{1, \ldots, \min\{m, n\}\}$, the problem

$$\min_{\mathbf{B} \in \mathbb{R}^{m \times n},\ \mathrm{rank}(\mathbf{B}) \leq k} \quad \|\mathbf{A} - \mathbf{B}\|_F^2$$

  has a solution given by $\mathbf{B}^\star = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$

- in this lecture, we will deal with the real matrices–the complex case follows along the same lines

Singular Value Decomposition

Theorem 7.1. Given any $\mathbf{A} \in \mathbb{R}^{m \times n}$, there exists a 3-tuple $(\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}) \in \mathbb{R}^{m \times m} \times \mathbb{R}^{m \times n} \times \mathbb{R}^{n \times n}$ such that

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T,$$

$\mathbf{U}$ and $\mathbf{V}$ are orthogonal, and $\mathbf{\Sigma}$ takes the form

$$[\mathbf{\Sigma}]_{ij} = \begin{cases} \sigma_i, & i = j \\ 0, & i \neq j \end{cases}, \qquad \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0, \; p = \min\{m, n\}.$$

- the above decomposition is called the singular value decomposition (SVD)
- σ_i is called the ith singular value
- $\mathbf{u}_i$ and $\mathbf{v}_i$ are called the ith left and right singular vectors, resp.

  $$\mathbf{u}_i^T \mathbf{A} = \sigma_i \mathbf{v}_i^T \Longleftarrow \mathbf{U}^T\mathbf{A} = \mathbf{\Sigma}\mathbf{V}^T \Longleftrightarrow \mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$$

  $$\Longleftrightarrow \mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma} \Longrightarrow \mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i \quad \text{for} \quad i = 1, \ldots, p$$

  $\mathbf{U}$ and $\mathbf{V}$ are called the left and right singular vector matrices, resp.
- the following notations may be used to denote singular values of a given $\mathbf{A}$

  $$\sigma_{\max}(\mathbf{A}) = \sigma_1(\mathbf{A}) \geq \sigma_2(\mathbf{A}) \geq \ldots \geq \sigma_p(\mathbf{A}) = \sigma_{\min}(\mathbf{A})$$

Different Ways of Writing out SVD

- partitioned form: let r be the number of nonzero singular values, and note $\sigma_1 \geq \ldots \geq \sigma_r > 0$, $\sigma_{r+1} = \ldots = \sigma_p = 0$. Then,

$$\mathbf{A} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{\Sigma}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix},$$

  where

  - $\tilde{\mathbf{\Sigma}} = \mathrm{Diag}(\sigma_1, \ldots, \sigma_r)$,
  - $\mathbf{U}_1 = [\ \mathbf{u}_1, \ldots, \mathbf{u}_r\ ] \in \mathbb{R}^{m \times r}$, $\mathbf{U}_2 = [\ \mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\ ] \in \mathbb{R}^{m \times (m-r)}$,
  - $\mathbf{V}_1 = [\ \mathbf{v}_1, \ldots, \mathbf{v}_r\ ] \in \mathbb{R}^{n \times r}$, $\mathbf{V}_2 = [\ \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\ ] \in \mathbb{R}^{n \times (n-r)}$.

- thin SVD (reduced SVD): $\mathbf{A} = \mathbf{U}_1 \tilde{\mathbf{\Sigma}} \mathbf{V}_1^T$

  - in contrast, the one in Theorem 7.1 is also called full SVD

- outer-product form (dyadic decomposition): $\mathbf{A} = \sum_{i=1}^{p} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$

SVD and Eigendecomposition

From the SVD $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, we see that

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{D}_1\mathbf{U}^T, \qquad \mathbf{D}_1 = \mathbf{\Sigma}\mathbf{\Sigma}^T = \mathrm{Diag}(\sigma_1^2, \ldots, \sigma_p^2, \underbrace{0, \ldots, 0}_{m-p \text{ zeros}}) \tag{$*$}$$

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{D}_2\mathbf{V}^T, \qquad \mathbf{D}_2 = \mathbf{\Sigma}^T\mathbf{\Sigma} = \mathrm{Diag}(\sigma_1^2, \ldots, \sigma_p^2, \underbrace{0, \ldots, 0}_{n-p \text{ zeros}}) \tag{$**$}$$

Observations:

- $(*)$ and $(**)$ are the SVD's of $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$, resp.
- $(*)$ and $(**)$ are the eigendecompositions of $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$, resp.
- the left singular matrix $\mathbf{U}$ of $\mathbf{A}$ is the eigenvector matrix of $\mathbf{A}\mathbf{A}^T$
- the right singular matrix $\mathbf{V}$ of $\mathbf{A}$ is the eigenvector matrix of $\mathbf{A}^T\mathbf{A}$
- the squares of nonzero singular values of $\mathbf{A}$, $\sigma_1^2, \ldots, \sigma_r^2$, are the nonzero eigenvalues of both $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$.
- the relation between SVD and eigendec. can be used for analysis and computations

Insights of the Proof of SVD

- the proof of SVD is constructive
- to see the insights, consider the special case of square nonsingular $\mathbf{A}$
- $\mathbf{A}\mathbf{A}^T$ is PD, and denote its eigendecomposition by

$$\mathbf{A}\mathbf{A}^T = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T, \qquad \text{with } \lambda_1 \geq \ldots \geq \lambda_n > 0.$$

- let $\boldsymbol{\Sigma} = \mathrm{Diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_m})$, $\mathbf{V} = \mathbf{A}^T\mathbf{U}\boldsymbol{\Sigma}^{-1}$
- it can be verified that $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T = \mathbf{A}$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}$

- how to prove the SVD in the general case? (requires a proof)

Uniqueness of SVD

- the singular values σ_i's are uniquely determined and the nonzero singular values are the positive square roots of the nonzero eigenvalues of $\mathbf{A}\mathbf{A}^T$ or, equivalently, of $\mathbf{A}^T\mathbf{A}$

- the multiplicity of a singular value σ of $\mathbf{A}$ is the multiplicity of σ^2 as an eigenvalue of $\mathbf{A}\mathbf{A}^T$ or, equivalently, of $\mathbf{A}^T\mathbf{A}$

- a singular value σ of $\mathbf{A}$ is simple (algebraic multiplicity is 1) if σ^2 is a simple eigenvalue of $\mathbf{A}\mathbf{A}^T$ or, equivalently, of $\mathbf{A}^T\mathbf{A}$

- uniqueness of SVD is higly related to the multiplicity of singular values and zero singular values of $\mathbf{A}$ and there are different kinds of characterizations; see Theorem 2.6.5 in **[Horn-Johnson'12]**.

Properties of SVD

Property 7.1. The following properties hold:

(a) $\mathbf{A}^T = \mathbf{V}\boldsymbol{\Sigma}^T\mathbf{U}^T$

(b) $\mathbf{A}$, $\mathbf{A}^*$, $\mathbf{A}^T$, and $\mathbf{A}^H$ have the same singular values

(c) $\mathbf{u}_i^T\mathbf{A}\mathbf{v}_i = \sigma_i$ for $i = 1, \ldots, p$, or, equivalently, in matrix form $\mathbf{U}^T\mathbf{A}\mathbf{V} = \boldsymbol{\Sigma}$

(d) $\mathrm{tr}(\mathbf{A}^T\mathbf{A}) = \mathrm{tr}(\mathbf{A}\mathbf{A}^T) = \sum_{i=1}^p \sigma_i^2$

(e) let $\mathbf{A} \in \mathbb{R}^{n\times n}$, $|\det(\mathbf{A})| = |\det(\boldsymbol{\Sigma})| = \prod_{i=1}^n \sigma_i$

(f) $\mathrm{rank}(\mathbf{A}) < q$ ($\mathbf{A}$ is singular) if and only if 0 is one singular value of $\mathbf{A}$

(g) let $\mathbf{A} \in \mathbb{S}^n$, the singular values are the absolute values of eigenvalues of $\mathbf{A}$

(h) if $\mathbf{A}$ is invertible, $\mathbf{A}^{-1} = \mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^T$ (can be used to compute matrix inversion)

(i) for othogonal $\mathbf{P}$ and $\mathbf{Q}$, SVD of $\mathbf{P}\mathbf{A}\mathbf{Q}^T$ is given by $\tilde{\mathbf{U}}\boldsymbol{\Sigma}\tilde{\mathbf{V}}^T$ where $\tilde{\mathbf{U}} = \mathbf{P}\mathbf{U}$ and $\tilde{\mathbf{V}} = \mathbf{Q}\mathbf{V}$, i.e., singular values are orthogonally invariant (i.e., $\sigma_i(\mathbf{A}) = \sigma_i(\mathbf{P}\mathbf{A}\mathbf{Q}^T)$) but singular vectors not

Properties of SVD

Property 7.2. The following properties hold:

(a) $\mathcal{R}(\mathbf{A}) = \mathcal{R}(\mathbf{U}_1)$, $\mathcal{R}(\mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^T) = \mathcal{R}(\mathbf{U}_2)$;

($\mathbf{U}_1$ and $\mathbf{U}_2$ forms a set of orthogonal bases for $\mathcal{R}(\mathbf{A})$ and $\mathcal{N}(\mathbf{A}^T)$ resp.)

(b) $\mathcal{R}(\mathbf{A}^T) = \mathcal{R}(\mathbf{V}_1)$, $\mathcal{R}(\mathbf{A}^T)^\perp = \mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{V}_2)$;

($\mathbf{V}_1$ and $\mathbf{V}_2$ forms a set of orthogonal bases for $\mathcal{R}(\mathbf{A}^T)$ and $\mathcal{N}(\mathbf{A})$ resp.)

(c) $\mathrm{rank}(\mathbf{A}) = r$ (the number of nonzero singular values).

Note:

- in practice, SVD can be used a numerical tool for computing bases of $\mathcal{R}(\mathbf{A})$, $\mathcal{N}(\mathbf{A}^T)$, $\mathcal{R}(\mathbf{A}^T)$, $\mathcal{N}(\mathbf{A})$
- we have previously learnt the following properties
  - $\mathrm{rank}(\mathbf{A}^T) = \mathrm{rank}(\mathbf{A})$
  - $\dim \mathcal{N}(\mathbf{A}) = n - \mathrm{rank}(\mathbf{A})$

  By SVD, the above properties are easily seen to be true
- SVD can also be used as a numerical tool to compute the rank of a matrix

Matrix Norms

- the definition of a norm of a matrix is the same as that of a vector:
- $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is a norm if

  (i) $f(\mathbf{A}) \geq 0$ for all $\mathbf{A}$;

  (ii) $f(\mathbf{A}) = 0$ if and only if $\mathbf{A} = \mathbf{0}$;

  (iii) $f(\mathbf{A}+\mathbf{B}) \leq f(\mathbf{A}) + f(\mathbf{B})$ for any $\mathbf{A}, \mathbf{B}$;

  (iv) $f(\alpha\mathbf{A}) = |\alpha| f(\mathbf{A})$ for any $\alpha, \mathbf{A}$

  (v) $f(\mathbf{AB}) \leq f(\mathbf{A}) f(\mathbf{B})$ for any $\mathbf{A}, \mathbf{B}$ (only for the case $m = n$)

"Elementwise" Norms

- "elementwise" norm: treat $\mathbf{A}$ as a $m \times n$ vector
- in general, for $p,\ q \geq 1$ it is given by

$$f(\mathbf{A}) = \left(\sum_{j=1}^{n} \left(\sum_{i=1}^{m} |a_{ij}|^p \right)^{\frac{q}{p}} \right)^{\frac{1}{q}}$$

- for $p = q = 2$, we have the Frobenius norm $\|\mathbf{A}\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2} = [\mathrm{tr}(\mathbf{A}^T\mathbf{A})]^{1/2}$
  - note Frobenius norm has the orthogonal invariance property, then $\|\mathbf{A}\|_F = \|\mathbf{U}^T\mathbf{A}\mathbf{V}\|_F = \|\mathbf{\Sigma}\|_F = \sqrt{\sigma_1^2 + \ldots + \sigma_r^2}$
- for $p = q = \infty$, we have the maximum norm $\|\mathbf{A}\|_\infty = \max_{i,j} |a_{ij}|$
- ...
- there are many other matrix norms

Induced Norms

- induced norm or operator norm: the function

$$f(\mathbf{A}) = \max_{\|\mathbf{x}\|_\beta \leq 1} \|\mathbf{A}\mathbf{x}\|_\alpha$$

  where $\|\cdot\|_\alpha, \|\cdot\|_\beta$ denote any vector norms, can be shown be to a norm

- induced p-norm: matrix norms induced by the vector p-norm $(p \geq 1)$

$$\|\mathbf{A}\|_p = \max_{\|\mathbf{x}\|_p \leq 1} \|\mathbf{A}\mathbf{x}\|_p$$

- it is known that
  - $\|\mathbf{A}\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|$
  - $\|\mathbf{A}\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$

- how about $p = 2$?

Induced Norms

- matrix 2-norm or spectral norm:

$$\|\mathbf{A}\|_2 = \sigma_{\max}(\mathbf{A}).$$

- proof:
  - for any $\mathbf{x}$ with $\|\mathbf{x}\|_2 \leq 1$,

$$\begin{aligned} \|\mathbf{A}\mathbf{x}\|_2^2 &= \|\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{x}\|_2^2 = \|\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{x}\|_2^2 \\ &\leq \sigma_1^2\|\mathbf{V}^T\mathbf{x}\|_2^2 = \sigma_1^2\|\mathbf{x}\|_2^2 \leq \sigma_1^2 \end{aligned}$$

  - $\|\mathbf{A}\mathbf{x}\|_2 = \sigma_1$ if we choose $\mathbf{x} = \mathbf{v}_1$

- implication to linear systems: let $\mathbf{y} = \mathbf{A}\mathbf{x}$ be a linear system. Under the input energy constraint $\|\mathbf{x}\|_2 \leq 1$, the system output energy $\|\mathbf{y}\|_2^2$ is maximized when $\mathbf{x}$ is chosen as the 1st right singular vector

- corollary: $\min_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2 = \sigma_{\min}(\mathbf{A})$ if $m \geq n$

- corollary: if $\mathbf{A}$ is invertible, $\|\mathbf{A}^{-1}\|_2 = 1/\sigma_{\min}(\mathbf{A})$

Induced Norms

Properties for the matrix 2-norm:

- $\|\mathbf{AB}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{B}\|_2$
  - in fact, $\|\mathbf{AB}\|_p \leq \|\mathbf{A}\|_p \|\mathbf{B}\|_p$ for any $p \geq 1$
- $\|\mathbf{Ax}\|_2 \leq \|\mathbf{A}\|_2 \|\mathbf{x}\|_2$
  - a special case of the 1st property
- $\|\mathbf{QAW}\|_2 = \|\mathbf{A}\|_2$ for any orthogonal $\mathbf{Q}, \mathbf{W}$
  - we also have $\|\mathbf{QAW}\|_F = \|\mathbf{A}\|_F$ for any orthogonal $\mathbf{Q}, \mathbf{W}$
- $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F \leq \sqrt{p}\|\mathbf{A}\|_2$ (here $p = \min\{m, n\}$)
  - proof: $\|\mathbf{A}\|_F = \|\mathbf{\Sigma}\|_F = \sqrt{\sum_{i=1}^{p} \sigma_i^2}$, and $\sigma_1^2 \leq \sum_{i=1}^{p} \sigma_i^2 \leq p\sigma_1^2$

Schatten Norms

- applying the p-norm to the vector of singular values of matrix $\mathbf{A}$

$$f(\mathbf{A}) = \left(\sum_{i=1}^{\min\{m,n\}} \sigma_i(\mathbf{A})^p \right)^{1/p}, \qquad p \geq 1,$$

  is known to be a norm and is called the Schatten p-norm

- Frobenius norm when $p = 2$; spectral norm when $p = \infty$
- nuclear norm (or trace norm) when $p = 1$:

$$\|\mathbf{A}\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i(\mathbf{A}) = \mathrm{tr}((\mathbf{A}^T\mathbf{A})^{\frac{1}{2}})$$

  - a special case of the Schatten p-norm
  - a way to prove that the nuclear norm is a norm:
    * show that $f(\mathbf{A}) = \max_{\|\mathbf{B}\|_2 \leq 1} \mathrm{tr}(\mathbf{B}^T\mathbf{A})$ is a norm
    * show that $f(\mathbf{A}) = \sum_{i=1}^{\min\{m,n\}} \sigma_i$
  - finds applications in rank approximation, e.g., for compressive sensing and matrix completion **[Recht-Fazel-Parrilo'10]**

Schatten Norms

- $\operatorname{rank}(\mathbf{A})$ is nonconvex in $\mathbf{A}$ and is arguably hard to do optimization with it
- Idea: the rank function can be expressed as

$$\operatorname{rank}(\mathbf{A}) = \sum_{i=1}^{\min\{m,n\}} \mathbb{1}\{\sigma_i(\mathbf{A}) \neq 0\},$$

  and why not approximate it by

$$f(\mathbf{A}) = \textstyle\sum_{i=1}^{\min\{m,n\}} \varphi(\sigma_i(\mathbf{A}))$$

  for some friendly function φ?

- nuclear norm

$$\|\mathbf{A}\|_* = \textstyle\sum_{i=1}^{\min\{m,n\}} \sigma_i(\mathbf{A})$$

  - uses $\varphi(z) = z$
  - is convex in $\mathbf{A}$
  - a convex envelope of $\operatorname{rank}(\mathbf{A})$

Linear Systems: Interpretation under SVD

- consider the linear system

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

  where $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the system matrix; $\mathbf{x} \in \mathbb{R}^n$ is the system input; $\mathbf{y} \in \mathbb{R}^m$ is the system output

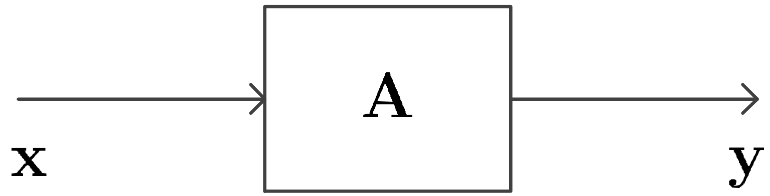
- by SVD we can write

$$\mathbf{y} = \mathbf{U}\tilde{\mathbf{y}}, \qquad \tilde{\mathbf{y}} = \boldsymbol{\Sigma}\tilde{\mathbf{x}}, \qquad \tilde{\mathbf{x}} = \mathbf{V}^T\mathbf{x}$$

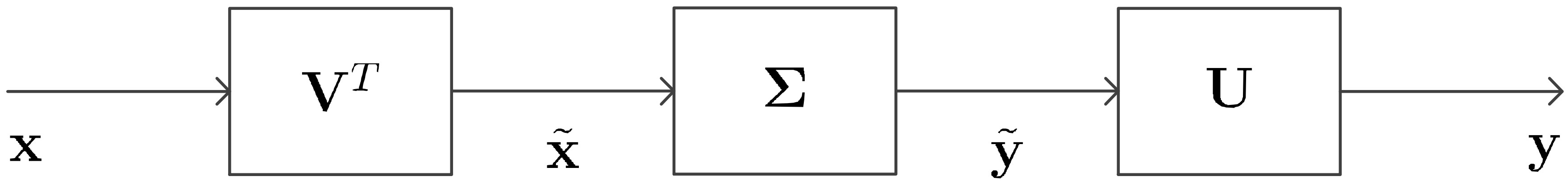
- **Implication:** *every* linear system $\mathbf{A}$ (a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$) works by performing three processes in cascade, namely,
  - rotate/reflect the system input $\mathbf{x}$ to form an intermediate system input $\tilde{\mathbf{x}}$
  - form an intermediate system output $\tilde{\mathbf{y}}$ by element-wise rescaling $\tilde{\mathbf{x}}$ w.r.t. σ_i's and by either removing some entires of $\tilde{\mathbf{x}}$ or adding some zeros
  - rotate/reflect $\tilde{\mathbf{y}}$ to form the system output $\mathbf{y}$

- **Implication:** every linear system $\mathbf{A}$ reduces to the diagonal matrix $\boldsymbol{\Sigma}$ when the range $\mathbf{y}$ is expressed in the basis of columns of $\mathbf{U}$ and the domain $\mathbf{x}$ is expressed in the basis of columns of $\mathbf{V}$

Linear Systems: Interpretation under SVD



(a) linear system

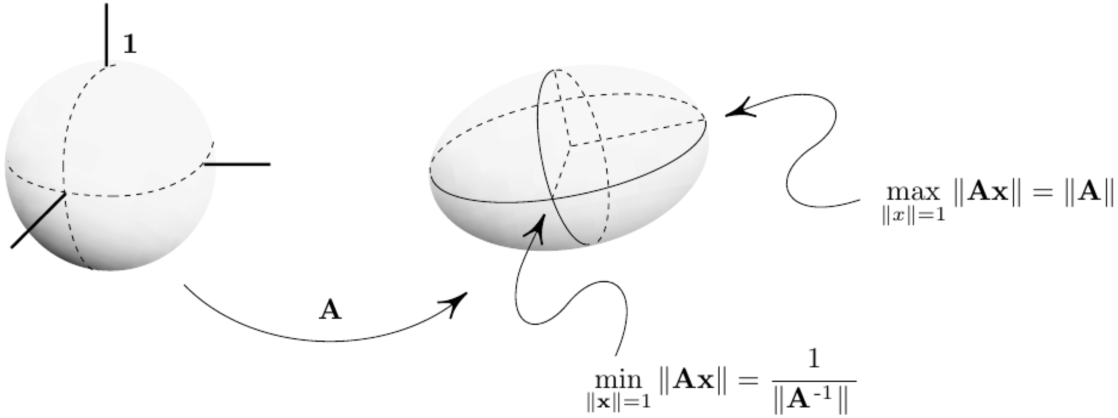


(b) equivalent system

Linear Systems: Interpretation under SVD

- SVD reveals the geometry about linear transformation $\mathbf{y} = \mathbf{A}\mathbf{x}$
- consider the transformation of a unit sphere in $\mathbb{R}^3$ under a nonsingular $\mathbf{A} \in \mathbb{R}^{3\times 3}$ and the singular values tell how much distortion can occur under $\mathbf{A}$

$$1 \geq \|\mathbf{x}\|_2^2 = \|\mathbf{A}^{-1}\mathbf{A}\mathbf{x}\|_2^2 = \|\mathbf{A}^{-1}\mathbf{y}\|_2^2 = \|\mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{U}^T\mathbf{y}\|_2^2 = \|\boldsymbol{\Sigma}^{-1}\mathbf{U}^T\mathbf{y}\|_2^2$$



(recall the result $\sigma_{\min}\|\mathbf{x}\|_2^2 \leq \|\mathbf{y}\|_2^2 = \|\mathbf{A}\mathbf{x}\|_2^2 \leq \sigma_{\max}\|\mathbf{x}\|_2^2$ for $m \geq n$)

- similar results apply to rectangular and singular $\mathbf{A}$
- **Fact:** the image of the unit sphere under *any* linear map $\mathbf{A}$ is a hyperellipse
- **Fact:** the amount of distortion of unit sphere under transformation $\mathbf{A}$ determines the degree to which uncertainties in a linear system $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be magnified

Linear Systems: Sensitivity Analysis

- Scenario:
  - let $\mathbf{A} \in \mathbb{R}^{n \times n}$ be nonsingular, and $\mathbf{y} \in \mathbb{R}^n$. Let $\mathbf{x}$ be the solution to

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

  - it is a well-determined linear system
  - consider a perturbed version of the above system: $\hat{\mathbf{A}} = \mathbf{A} + \Delta\mathbf{A}, \hat{\mathbf{y}} = \mathbf{y} + \Delta\mathbf{y}$, where $\Delta\mathbf{A}$ and $\Delta\mathbf{y}$ are errors. Let $\hat{\mathbf{x}}$ be a solution to the perturbed system

$$\hat{\mathbf{y}} = \hat{\mathbf{A}}\hat{\mathbf{x}}.$$

- Problem: analyze how the solution error $\|\hat{\mathbf{x}} - \mathbf{x}\|_2$ scales with $\Delta\mathbf{A}$ and $\Delta\mathbf{y}$

- remark: $\Delta\mathbf{A}$ and $\Delta\mathbf{y}$ may be floating point errors, measurement errors, etc

Linear Systems: Sensitivity Analysis

- the condition number of a given nonsingular matrix $\mathbf{A}$ is defined as

$$\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$$

  - $\kappa(\mathbf{A}) \geq 1$
  - $\mathbf{A}$ is said to be well-conditioned if $\kappa(\mathbf{A})$ is small
  - $\mathbf{A}$ is said to be ill-conditioned if $\kappa(\mathbf{A})$ is very large; that refers to cases where $\mathbf{A}$ is close to singular (high linear dependence between columns or rows of $\mathbf{A}$)
  - it is customary to denote $\kappa(\mathbf{A}) = \infty$ if $\mathbf{A}$ is a singular matrix

- the 2-norm condition number of a given nonsingular matrix $\mathbf{A}$ is given by

$$\kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2\|\mathbf{A}^{-1}\|_2 = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})}$$

- $\kappa_2(\mathbf{A}) = 1$ if $\mathbf{A}$ is a multiple of an orthogonal matrix (perfectly conditioned)
- if not specially specified, the condition number is commonly referred to as $\kappa_2(\mathbf{A})$

Linear Systems: Sensitivity Analysis

Theorem 7.2. If $\mathbf{A}$ is known exactly and there is an uncertainty $\Delta\mathbf{y}$, then

$$\kappa_2^{-1}(\mathbf{A})\frac{\|\Delta\mathbf{y}\|_2}{\|\mathbf{y}\|_2} \le \frac{\|\hat{\mathbf{x}}-\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le \kappa_2(\mathbf{A})\frac{\|\Delta\mathbf{y}\|_2}{\|\mathbf{y}\|_2}.$$

(requires a proof)

- if $\mathbf{A}$ is well-conditioned, a small uncertainty in $\mathbf{y}$ cannot produce a very large solution error
- if $\mathbf{A}$ is ill-conditioned, a small uncertainty in $\mathbf{y}$ can produce a very large solution error; or a large uncertainty in $\mathbf{y}$ can produce a very small solution error, which depends on the "direction" of $\Delta\mathbf{y}$

Theorem 7.3. If $\mathbf{y}$ is known exactly and there is an uncertainty $\Delta\mathbf{A}$, then

$$\frac{\|\hat{\mathbf{x}}-\mathbf{x}\|_2}{\|\hat{\mathbf{x}}\|_2} \le \kappa_2(\mathbf{A})\frac{\|\Delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2} \quad \text{and} \quad \frac{\|\hat{\mathbf{x}}-\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le \frac{1}{1-\kappa_2(\mathbf{A})\frac{\|\Delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2}}\kappa_2(\mathbf{A})\frac{\|\Delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2}.$$

(proof by yourself)

Linear Systems: Sensitivity Analysis

Theorem 7.4. If there are uncertainties $\Delta\mathbf{A}$ and $\Delta\mathbf{y}$, then

$$\frac{\|\hat{\mathbf{x}}-\mathbf{x}\|_2}{\|\hat{\mathbf{x}}\|_2} \leq \kappa_2(\mathbf{A})\left(\frac{\|\Delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2}+\frac{\|\Delta\mathbf{y}\|_2}{\|\mathbf{A}\|_2\|\hat{\mathbf{x}}\|_2}\right)$$

and

or $$\frac{\|\hat{\mathbf{x}}-\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \frac{1}{1-\kappa_2(\mathbf{A})\frac{\|\Delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2}}\kappa_2(\mathbf{A})\left(\frac{\|\Delta\mathbf{A}\|_2}{\|\mathbf{A}\|_2}+\frac{\|\Delta\mathbf{y}\|_2}{\|\mathbf{y}\|_2}\right).$$

(proof by yourself)

Linear Systems: Sensitivity Analysis

Theorem 7.5. Let $\varepsilon > 0$ be a constant such that

$$\frac{\|\Delta \mathbf{A}\|_2}{\|\mathbf{A}\|_2} \le \varepsilon, \qquad \frac{\|\Delta \mathbf{y}\|_2}{\|\mathbf{y}\|_2} \le \varepsilon.$$

If ε is sufficiently small such that $\varepsilon\kappa_2(\mathbf{A}) < 1$, then

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le \frac{2\varepsilon\kappa_2(\mathbf{A})}{1 - \varepsilon\kappa_2(\mathbf{A})}.$$

(requires a proof)

- Implications:
  - for small errors and in the worst-case sense, the relative error $\|\hat{\mathbf{x}} - \mathbf{x}\|_2 / \|\mathbf{x}\|_2$ tends to increase with the condition number
  - in particular, for $\varepsilon\kappa_2(\mathbf{A}) \le \frac{1}{2}$, the error bound can be simplified to

    $$\frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \le 4\varepsilon\kappa_2(\mathbf{A})$$

    where the error bound scales linearly with the condition number

Linear Systems: Sensitivity Analysis

- Scenario:
  - let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be nonsingular, and $\mathbf{y} \in \mathbb{R}^m$. A vector $\mathbf{x}_{\mathsf{LS}}$ is an optimal solution to the LS problem
  $$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$$
  if and only if it satisfies the normal equation
  $$\mathbf{A}^T\mathbf{A}\mathbf{x}_{\mathsf{LS}} = \mathbf{A}^T\mathbf{y}.$$
  - consider a perturbed version of the above system: $\hat{\mathbf{A}} = \mathbf{A} + \Delta\mathbf{A}, \hat{\mathbf{y}} = \mathbf{y} + \Delta\mathbf{y}$, where $\Delta\mathbf{A}$ and $\Delta\mathbf{y}$ are errors. Let $\hat{\mathbf{x}}$ be a solution to the perturbed system
  $$\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{x}}_{\mathsf{LS}} = \hat{\mathbf{A}}^T\hat{\mathbf{y}}.$$
- Problem: analyze how the solution error $\|\hat{\mathbf{x}} - \mathbf{x}\|_2$ scales with $\Delta\mathbf{A}$ and $\Delta\mathbf{y}$

Linear Systems: Sensitivity Analysis

- note that the condition number

$$\kappa_2(\mathbf{A}^T\mathbf{A}) = (\kappa_2(\mathbf{A}))^2$$

- implication: we should avoid directly solving the normal equation

- when the QR decompostion $\mathbf{A} = \mathbf{QR}$ is applied for LS solving, we have

$$\kappa_2(\mathbf{Q}) = 1 \quad \text{and} \quad \kappa_2(\mathbf{A}) = \kappa_2(\mathbf{Q}^T\mathbf{A}) = \kappa_2(\mathbf{R})$$

  in which case the influence of $\Delta\mathbf{A}$ and $\Delta\mathbf{y}$ to the solution error in LS is propotional to $\kappa_2(\mathbf{A})$ in the same way as in the linear system

- implication: LS via QR is more numerically stable

- Question: how to tackle the ill-conditioned $\mathbf{A}$? one solution is the total least squares method (in Topic 8: Least Squares Revisited) which relies on the SVD

Linear Systems: Solution via SVD

- Problem: given *general* $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{y} \in \mathbb{R}^m$, determine
  - whether $\mathbf{y} = \mathbf{A}\mathbf{x}$ has a solution
  - what is the solution

- by SVD it can be shown that

$$
\begin{aligned}
\mathbf{y} = \mathbf{A}\mathbf{x} \quad &\Longleftrightarrow \quad \mathbf{y} = \mathbf{U}_1 \tilde{\boldsymbol{\Sigma}} \mathbf{V}_1^T \mathbf{x} \\
&\Longleftrightarrow \quad \mathbf{U}_1^T \mathbf{y} = \tilde{\boldsymbol{\Sigma}} \mathbf{V}_1^T \mathbf{x}, \ \mathbf{U}_2^T \mathbf{y} = \mathbf{0} \\
&\Longleftrightarrow \quad \mathbf{V}_1^T \mathbf{x} = \tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{U}_1^T \mathbf{y}, \ \mathbf{U}_2^T \mathbf{y} = \mathbf{0} \\
&\Longleftrightarrow \quad \begin{array}{l} \mathbf{x} = \mathbf{V}_1 \tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{U}_1^T \mathbf{y} + \boldsymbol{\eta}, \ \text{for any } \boldsymbol{\eta} \in \mathcal{R}(\mathbf{V}_2) = \mathcal{N}(\mathbf{A}), \\ \mathbf{U}_2^T \mathbf{y} = \mathbf{0} \end{array}
\end{aligned}
$$

- the linear system $\mathbf{y} = \mathbf{A}\mathbf{x}$ is said to be consistent if $\mathbf{U}_2^T \mathbf{y} = \mathbf{0}$

Linear Systems: Solution via SVD

- let us consider specific cases of the linear system solution characterization

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad \Longleftrightarrow \quad \begin{array}{l} \mathbf{x} = \mathbf{V}_1 \tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{U}_1^T \mathbf{y} + \boldsymbol{\eta}, \text{ for any } \boldsymbol{\eta} \in \mathcal{R}(\mathbf{V}_2) = \mathcal{N}(\mathbf{A}), \\ \mathbf{U}_2^T \mathbf{y} = \mathbf{0} \end{array}$$

- Case (a): full-column rank $\mathbf{A}$, i.e., $r = n \leq m$
  - there is no $\mathbf{V}_2$, and $\mathbf{U}_2^T \mathbf{y} = \mathbf{0}$ is equivalent to $\mathbf{y} \in \mathcal{R}(\mathbf{U}_1) = \mathcal{R}(\mathbf{A})$
  - Result: the linear system has a solution if and only if $\mathbf{y} \in \mathcal{R}(\mathbf{A})$, and the solution, if exists, is uniquely given by $\mathbf{x} = \mathbf{V} \tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{U}_1^T \mathbf{y}$

- Case (b): full-row rank $\mathbf{A}$, i.e., $r = m \leq n$
  - there is no $\mathbf{U}_2$
  - Result: the linear system always has a solution, and the solution is given by $\mathbf{x} = \mathbf{V}_1 \tilde{\boldsymbol{\Sigma}}^{-1} \mathbf{U}^T \mathbf{y} + \boldsymbol{\eta}$ for any $\boldsymbol{\eta} \in \mathcal{N}(\mathbf{A})$

Least Squares: Solution via SVD

- consider the LS problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2$$

  for *general* $\mathbf{A}\in\mathbb{R}^{m\times n}$

- we have, for any $\mathbf{x}\in\mathbb{R}^n$,

$$\begin{aligned}
\|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2^2 = \|\mathbf{y}-\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{x}\|_2^2 &= \|\mathbf{U}^T\mathbf{y}-\boldsymbol{\Sigma}\mathbf{V}^T\mathbf{x}\|_2^2 \\
&= \left\| \begin{bmatrix}\mathbf{U}_1^T\\ \mathbf{U}_2^T\end{bmatrix}\mathbf{y} - \begin{bmatrix}\tilde{\boldsymbol{\Sigma}}\mathbf{V}_1^T\\ \mathbf{0}\end{bmatrix}\mathbf{x}\right\|_2^2 \\
&= \|\mathbf{U}_1^T\mathbf{y}-\tilde{\boldsymbol{\Sigma}}\mathbf{V}_1^T\mathbf{x}\|_2^2 + \|\mathbf{U}_2^T\mathbf{y}\|_2^2 \\
&\geq \|\mathbf{U}_2^T\mathbf{y}\|_2^2
\end{aligned}$$

- the equality above is attained if $\mathbf{x}$ satisfies $\mathbf{U}_1^T\mathbf{y}=\tilde{\boldsymbol{\Sigma}}\mathbf{V}_1^T\mathbf{x}$, and that leads to an LS solution

$$\begin{aligned}
\mathbf{U}_1^T\mathbf{y}=\tilde{\boldsymbol{\Sigma}}\mathbf{V}_1^T\mathbf{x} \quad &\Longleftrightarrow \quad \mathbf{V}_1^T\mathbf{x}=\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{U}_1^T\mathbf{y} \\
&\Longleftrightarrow \quad \mathbf{x}=\mathbf{V}_1\tilde{\boldsymbol{\Sigma}}^{-1}\mathbf{U}_1^T\mathbf{y}+\boldsymbol{\eta},\ \text{for any } \boldsymbol{\eta}\in\mathcal{R}(\mathbf{V}_2)=\mathcal{N}(\mathbf{A})
\end{aligned}$$

Pseudo-Inverse

The pseudo-inverse (or Moore-Penrose inverse) of a matrix $\mathbf{A}$ is defined as

$$\mathbf{A}^{\dagger} = \mathbf{V}_1 \tilde{\mathbf{\Sigma}}^{-1} \mathbf{U}_1^T \in \mathbb{R}^{n \times m}.$$

From the above definition, we can show that

- let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{A}^{\dagger}$ always exists and unique
- $\mathbf{x}_{\mathsf{LS}} = \mathbf{A}^{\dagger}\mathbf{y} + \boldsymbol{\eta}$ for any $\boldsymbol{\eta} \in \mathcal{R}(\mathbf{V}_2)$; the same applies to linear sys. $\mathbf{y} = \mathbf{A}\mathbf{x}$
- it can be easily shown that

$$\mathbf{A}^{\dagger} = \mathbf{V}\mathbf{\Sigma}^{\dagger}\mathbf{U}^T \quad \text{with} \quad \mathbf{\Sigma}^{\dagger} = \begin{bmatrix} \tilde{\mathbf{\Sigma}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

- we also have $\mathbf{A}^{\dagger} = \mathbf{V}_1 \tilde{\mathbf{\Sigma}}^{-1} \mathbf{U}_1^T = \sum_{i=1}^{p} \frac{1}{\sigma_i} \mathbf{v}_i \mathbf{u}_i^T = \sum_{i=1}^{r} \frac{1}{\sigma_i} \mathbf{v}_i \mathbf{u}_i^T$

Pseudo-Inverse

- $\mathbf{A}^\dagger$ satisfies the Moore-Penrose conditions: (i) $\mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}$; (ii) $\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger$; (iii) $\mathbf{A}\mathbf{A}^\dagger$ is symmetric; (iv) $\mathbf{A}^\dagger\mathbf{A}$ is symmetric

- note: in general, $\mathbf{A}\mathbf{A}^\dagger \neq \mathbf{I}$ and $\mathbf{A}^\dagger\mathbf{A} \neq \mathbf{I}$

some properties of the Pseudo-Inverse:

- $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$
- $(\mathbf{A}^T)^\dagger = (\mathbf{A}^\dagger)^T$, $(\mathbf{A}^H)^\dagger = (\mathbf{A}^\dagger)^H$, $(\mathbf{A}^*)^\dagger = (\mathbf{A}^\dagger)^*$
- $(a\mathbf{A}^\dagger) = a^{-1}(\mathbf{A})^\dagger$ for $a \neq 0$
- $\mathrm{rank}(\mathbf{A}^\dagger) = \mathrm{rank}(\mathbf{A}) = \mathrm{rank}(\mathbf{A}^\dagger\mathbf{A}) = \mathrm{rank}(\mathbf{A}\mathbf{A}^\dagger)$
- $(\mathbf{A}\mathbf{A}^T)^\dagger = (\mathbf{A}^T)^\dagger(\mathbf{A})^\dagger$, $(\mathbf{A}^T\mathbf{A})^\dagger = (\mathbf{A})^\dagger(\mathbf{A}^T)^\dagger$
- $(\mathbf{A}\mathbf{A}^T)^\dagger\mathbf{A}\mathbf{A}^T = \mathbf{A}\mathbf{A}^\dagger$, $(\mathbf{A}^T\mathbf{A})^\dagger\mathbf{A}^T\mathbf{A} = \mathbf{A}^\dagger\mathbf{A}$
- for orthogonal $\mathbf{P}$, $\mathbf{Q}$, $(\mathbf{P}\mathbf{A}\mathbf{Q})^\dagger = \mathbf{Q}^T\mathbf{A}^\dagger\mathbf{P}^T$

Pseudo-Inverse

some properties of the pseudo-inverse:

- $\mathbf{A}^\dagger = (\mathbf{A}^T\mathbf{A})^\dagger\mathbf{A}^T = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^\dagger$
- specially, when $\mathbf{A}$ has full-column rank
  - the pseudo-inverse also equals $\mathbf{A}^\dagger = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$
  - $\mathbf{A}^\dagger\mathbf{A} = \mathbf{I}$ (hence called left inverse in this case)
- specially, when $\mathbf{A}$ has full-row rank
  - the pseudo-inverse also equals $\mathbf{A}^\dagger = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$
  - $\mathbf{A}\mathbf{A}^\dagger = \mathbf{I}$ (hence called right inverse in this case)
- specially, when $\mathbf{A}$ is square and has full rank
  - the pseudo-inverse also equals $\mathbf{A}^\dagger = \mathbf{A}^{-1}$

- note: for $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$, in general (a) $(\mathbf{A}\mathbf{B})^\dagger \neq \mathbf{B}^\dagger\mathbf{A}^\dagger$; (b) $\mathbf{A}\mathbf{A}^\dagger \neq \mathbf{A}^\dagger\mathbf{A}$; (c) $(\mathbf{A}^k)^\dagger \neq (\mathbf{A}^\dagger)^k$; (d) positive eigenvalues of $\mathbf{A}^\dagger$ are not reciprocals of those of $\mathbf{A}$

Computation of the Pseudo-Inverse

- computation via SVD
  - reply on the computation of the SVD

- computation via QR decomposition (possibly with column pivoting)
  - for example, let $\mathbf{A} \in \mathbb{R}^{m\times n}$ with full column rank and the thin QR is given by $\mathbf{A} = \mathbf{Q}_1\mathbf{R}_1$, then
$$\mathbf{A}^\dagger = \mathbf{R}_1^{-1}\mathbf{Q}_1^T$$

Orthogonal Projections

- with SVD, the orthogonal projections of $\mathbf{y}$ onto $\mathcal{R}(\mathbf{A})$ and $\mathcal{R}(\mathbf{A})^\perp$ are, resp.,

$$\Pi_{\mathcal{R}(\mathbf{A})}(\mathbf{y}) = \mathbf{A}\mathbf{x}_{\mathsf{LS}} = \mathbf{A}\mathbf{A}^\dagger\mathbf{y} = \mathbf{U}_1\mathbf{U}_1^T\mathbf{y}$$

$$\Pi_{\mathcal{R}(\mathbf{A})^\perp}(\mathbf{y}) = \mathbf{y} - \mathbf{A}\mathbf{x}_{\mathsf{LS}} = (\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger)\mathbf{y} = \mathbf{U}_2\mathbf{U}_2^T\mathbf{y}$$

- the orthogonal projector (projection matrix) and orthogonal complement projector of $\mathbf{A}$ are resp. defined as

$$\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{A}^\dagger = \mathbf{U}_1\mathbf{U}_1^T, \qquad \mathbf{P}_\mathbf{A}^\perp = (\mathbf{I} - \mathbf{A}\mathbf{A}^\dagger) = \mathbf{U}_2\mathbf{U}_2^T$$

- properties (easy to show):
  - $\mathbf{P}_\mathbf{A}$ is idempotent, i.e., $\mathbf{P}_\mathbf{A}^2 = \mathbf{P}_\mathbf{A}\mathbf{P}_\mathbf{A} = \mathbf{P}_\mathbf{A}$
  - $\mathbf{P}_\mathbf{A}$ is symmetric
  - the eigenvalues of $\mathbf{P}_\mathbf{A}$ are either 0 or 1
  - $\mathcal{R}(\mathbf{P}_\mathbf{A}) = \mathcal{R}(\mathbf{A})$
  - the same properties above apply to $\mathbf{P}_\mathbf{A}^\perp$, and $\mathbf{I} = \mathbf{P}_\mathbf{A} + \mathbf{P}_\mathbf{A}^\perp$

Orthogonal Projections

- similarly, the orthogonal projector (projection matrix) and orthogonal complement projector of $\mathbf{A}^T$ are resp. defined as

$$\mathbf{P}_{\mathbf{A}^T} = \mathbf{A}^\dagger \mathbf{A} = \mathbf{V}_1 \mathbf{V}_1^T = \mathbf{P}_{\mathbf{A}^\dagger}, \qquad \mathbf{P}_{\mathbf{A}^T}^\perp = (\mathbf{I} - \mathbf{A}^\dagger \mathbf{A}) = \mathbf{V}_2 \mathbf{V}_2^T = \mathbf{P}_{\mathbf{A}^\dagger}^\perp$$

- $\mathbf{P}_{\mathbf{A}^T}$ and $\mathbf{P}_{\mathbf{A}^T}^\perp$ are the orthogonal projections onto $\mathcal{R}(\mathbf{A}^T)$ (or $\mathcal{R}(\mathbf{A}^\dagger)$) and $\mathcal{R}(\mathbf{A}^T)^\perp$ (or $\mathcal{R}(\mathbf{A}^\dagger)^\perp$) resp.

we also have the following properties:

- $\mathcal{R}(\mathbf{A}\mathbf{A}^\dagger) = \mathcal{R}(\mathbf{A}\mathbf{A}^T) = \mathcal{R}(\mathbf{A})$
- $\mathcal{R}(\mathbf{A}^\dagger \mathbf{A}) = \mathcal{R}(\mathbf{A}^T \mathbf{A}) = \mathcal{R}(\mathbf{A}^T) = \mathcal{R}(\mathbf{A}^\dagger)$
- $\mathcal{N}(\mathbf{A}\mathbf{A}^\dagger) = \mathcal{N}(\mathbf{A}\mathbf{A}^T) = \mathcal{N}(\mathbf{A}^T) = \mathcal{N}(\mathbf{A}^\dagger)$
- $\mathcal{N}(\mathbf{A}^\dagger \mathbf{A}) = \mathcal{N}(\mathbf{A}^T \mathbf{A}) = \mathcal{N}(\mathbf{A})$

Minimum 2-Norm Solution to Underdetermined Linear Systems

- consider solving the linear system $\mathbf{y} = \mathbf{A}\mathbf{x}$ when $\mathbf{A}$ is fat
- this is an underdetermined problem: we have more unknowns n than the number of equations m
- assume that $\mathbf{A}$ has full row rank. By now we know that any

$$\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{y} + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \in \mathcal{R}(\mathbf{V}_2)$$

  is a solution to $\mathbf{y} = \mathbf{A}\mathbf{x}$, but we may want to grab one solution only

- Idea: discard $\boldsymbol{\eta}$ and take $\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{y}$ as our solution
- Question: does discarding $\boldsymbol{\eta}$ make sense?
- Answer: it makes sense under the minimum 2-norm problem formulation

$$\min_{\mathbf{x} \in \mathbb{R}^n} \ \|\mathbf{x}\|_2^2 \qquad \text{s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}$$

  It can be shown that the solution is uniquely given by $\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{y}$ (try the proof)

Minimum 2-Norm Solution to Linear System and Least Squares

generally, for any $\mathbf{A}$ and $\mathbf{y}$

- when $\mathbf{y} = \mathbf{A}\mathbf{x}$ is consistent, $\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{y}$ is the unique (linear system/least squares) solution of minimum 2-norm
- when $\mathbf{y} = \mathbf{A}\mathbf{x}$ is inconsistent, $\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{y}$ is the unique least squares solution of minimum 2-norm
- specifically, when $\mathbf{A}$ is full-colum rank, $\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{y}$ is the unique solution

Generalized Condition Number

- the condition number of a general matrix $\mathbf{A}$ is defined as

$$\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{\dagger}\|$$

- Scenario:
  - let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be a general matrix, and $\mathbf{y} \in \mathbb{R}^{n}$. Let $\mathbf{x}$ be the minimum 2-norm solution to
$$\mathbf{y} = \mathbf{A}\mathbf{x}.$$
  - consider a perturbed version of the above system: $\hat{\mathbf{y}} = \mathbf{y} + \Delta\mathbf{y}$, where $\Delta\mathbf{y}$ is the error. Let $\Delta\mathbf{x} = \hat{\mathbf{x}} - \mathbf{x}$ be the minimum 2-norm solution to
$$\Delta\mathbf{y} = \mathbf{A}\Delta\mathbf{x}.$$

Theorem 7.6. If $\mathbf{A}$ is known exactly and there is an uncertainty $\Delta\mathbf{y}$, then

$$\kappa_2^{-1}(\mathbf{A})\frac{\|\Delta\mathbf{y}\|_2}{\|\mathbf{y}\|_2} \leq \frac{\|\hat{\mathbf{x}} - \mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \kappa_2(\mathbf{A})\frac{\|\Delta\mathbf{y}\|_2}{\|\mathbf{y}\|_2}.$$

- similar results hold for other scenarios...

Low-Rank Matrix Approximation

Aim: given a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and an integer k with $0 \leq k \leq \operatorname{rank}(\mathbf{A})$, find a matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$ such that $\operatorname{rank}(\mathbf{B}) \leq k$ and $\mathbf{B}$ best approximates $\mathbf{A}$

- it is somehow unclear about what a "best approximation" means, and we will specify one later
- closely related to the matrix factorization problem considered in Topic 3: Least Squares
- applications: PCA, dimensionality reduction,...—the same kind of applications in matrix factorization
- truncated SVD: denote

$$\mathbf{A}_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$$

  where the kth "partial sum" captures as much of the energy of $\mathbf{A}$ as possible, and the meaning of "energy" will be specified later
- then perform the aforementioned approximation by choosing $\mathbf{B} = \mathbf{A}_k$

Toy Application Example: Image Compression

- let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be a matrix whose (i, j)th entry a_{ij} stores the (i, j)th pixel of an image

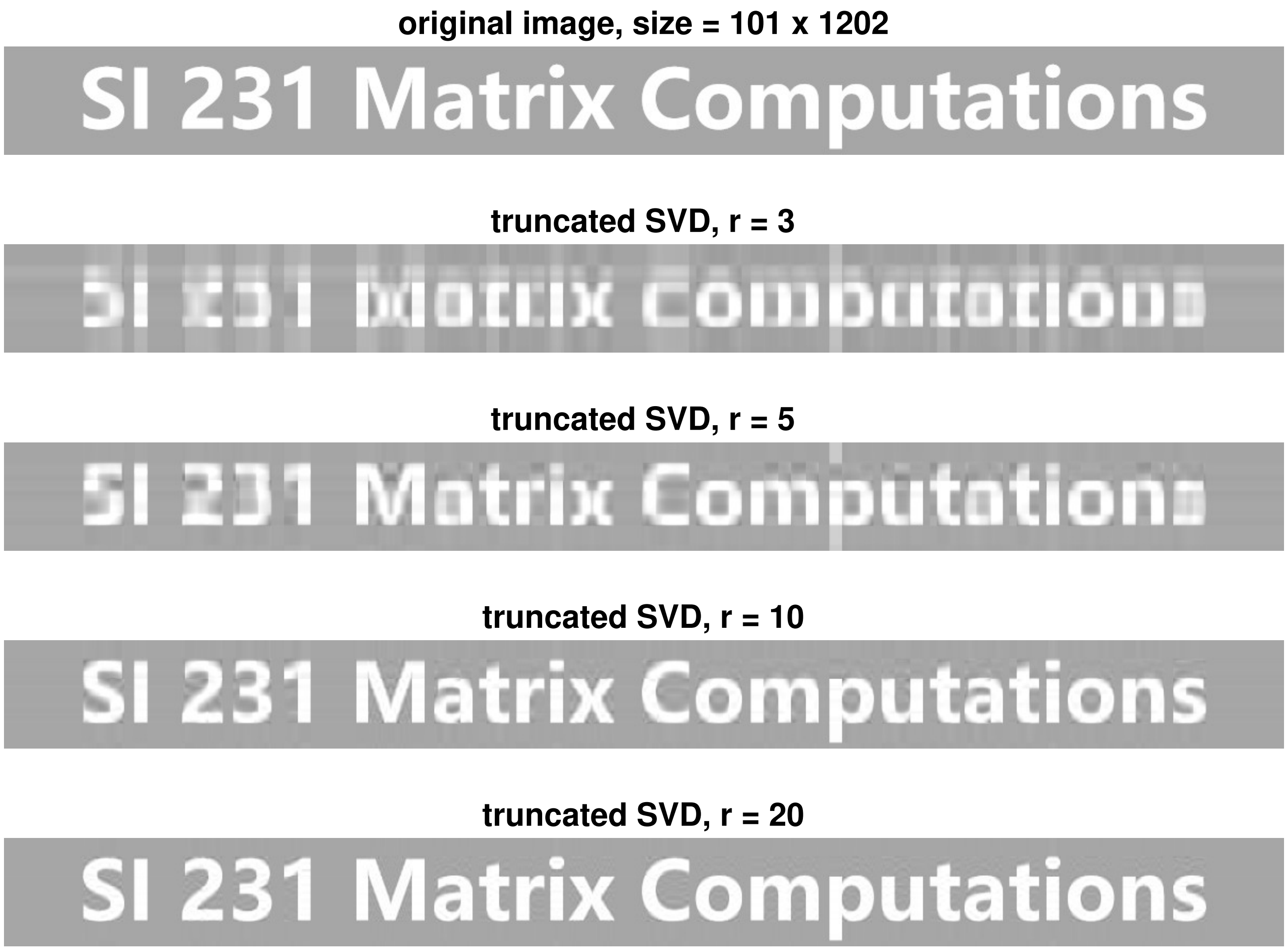
- memory size for storing $\mathbf{A}$: mn

- truncated SVD: store $\{\mathbf{u}_i, \sigma_i \mathbf{v}_i\}_{i=1}^{k}$ instead of the full $\mathbf{A}$, and recover the image by $\mathbf{B} = \mathbf{A}_k$

- memory size for truncated SVD: $(m + n)k$

  – much less than mn if $k \ll \min\{m, n\}$

Toy Application Example: Image Compression



Low-Rank Matrix Approximation

- truncated SVD provides the best approximation in the LS sense:
  Theorem 7.7 (Eckart-Young-Mirsky)**.** Consider the following problem

  $$\min_{\mathbf{B}\in\mathbb{R}^{m\times n},\ \mathrm{rank}(\mathbf{B})\leq k} \|\mathbf{A}-\mathbf{B}\|_F^2$$

  where $\mathbf{A}\in\mathbb{R}^{m\times n}$ and $k\in\{1,\ldots,p\}$ are given. The truncated SVD $\mathbf{A}_k$ is an optimal solution to the above problem and the minimum is $\sum_{i=k+1}^{p}\sigma_i^2$
  (a proof is given later)
- also note the matrix 2-norm version of the Eckart-Young-Mirsky theorem:
  Theorem 7.8. Consider the following problem

  $$\min_{\mathbf{B}\in\mathbb{R}^{m\times n},\ \mathrm{rank}(\mathbf{B})\leq k} \|\mathbf{A}-\mathbf{B}\|_2^2$$

  where $\mathbf{A}\in\mathbb{R}^{m\times n}$ and $k\in\{1,\ldots,p\}$ are given. The truncated SVD $\mathbf{A}_k$ is an optimal solution to the above problem and the minimum is σ_{k+1}^2
  (cf. Theorem 2.4.8 in **[Golub-Van Loan'13]**)
- the energy mentioned before is defined by either the Frobenius norm or the 2-norm

Low-Rank Matrix Approximation

- recall the matrix factorization problem in Topic 3:

$$\min_{\mathbf{A}\in\mathbb{R}^{m\times k},\mathbf{B}\in\mathbb{R}^{k\times n}} \|\mathbf{Y}-\mathbf{AB}\|_F^2$$

  where $k \leq \min\{m,n\}$; $\mathbf{A}$ denotes a basis matrix; $\mathbf{B}$ is the coefficient matrix

- the matrix factorizaton problem may be reformulated as (verify)

$$\min_{\mathbf{Z}\in\mathbb{R}^{m\times n},\mathrm{rank}(\mathbf{Z})\leq k} \|\mathbf{Y}-\mathbf{Z}\|_F^2,$$

  and the truncated SVD $\mathbf{Y}_k = \sum_{i=1}^{k}\sigma_i\mathbf{u}_i\mathbf{v}_i^T$, where $\mathbf{Y}=\mathbf{U\Sigma V}^T$ denotes the SVD of $\mathbf{Y}$, is an optimal solution by Theorem 7.7

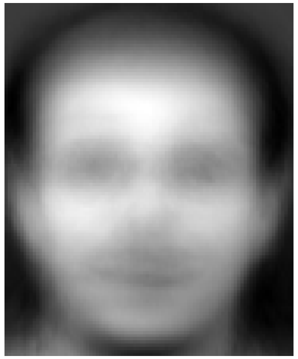
- thus, an optimal solution to the matrix factorization problem is

$$\mathbf{A} = [\ \mathbf{u}_1,\ldots,\mathbf{u}_k\ ], \qquad \mathbf{B} = [\ \sigma_1\mathbf{v}_1,\ldots,\sigma_k\mathbf{v}_k\ ]^T$$

Toy Demo: Dimensionality Reduction of a Face Image Dataset

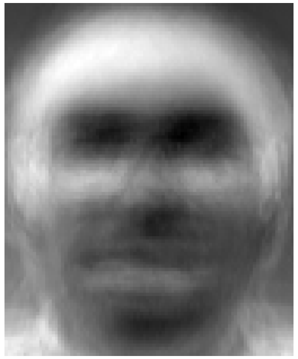
A face image dataset. Image size $= 112 \times 92$, number of face images $= 400$. Each $\mathbf{x}_i$ is the vectorization of one face image, leading to $m = 112 \times 92 = 10304$, $n = 400$.

Toy Demo: Dimensionality Reduction of a Face Image Dataset

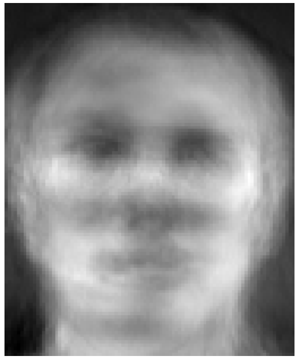


Mean face

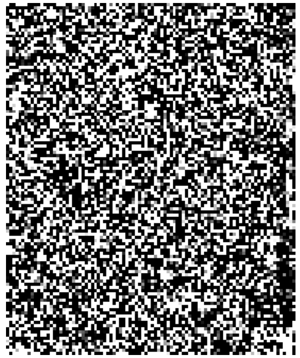
1st principal left singular vector



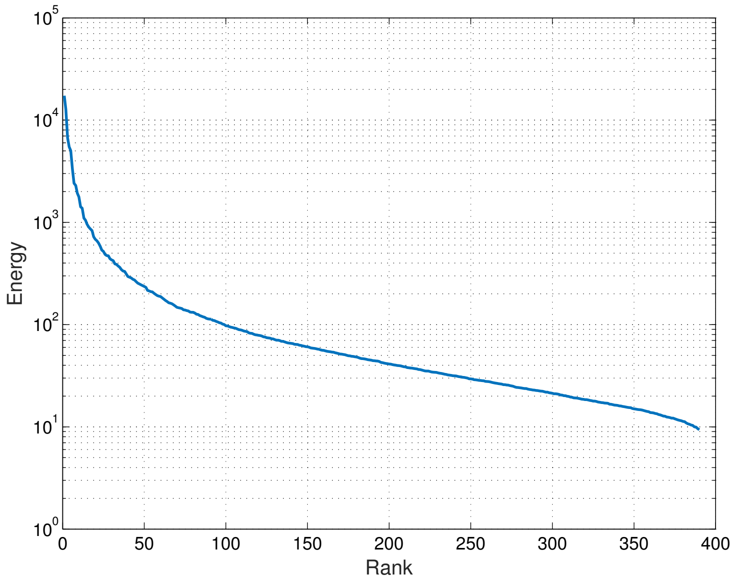
2nd principal left singular vector



3rd principal left singular vector



400th left singular vector



Energy Concentration

Variational Characterizations and Singular Value Inequalities

Similar to variational characterization of eigenvalues of Hermitian & real symmetric matrices in Topic 5, we can derive various variational characterization results for singular values, e.g.,

- Courant-Fischer characterization:

$$\sigma_k(\mathbf{A}) = \min_{\dim \mathcal{S}_{n-k+1} \subseteq \mathbb{R}^n} \max_{\mathbf{x} \in \mathcal{S}_{n-k+1},\ \|\mathbf{x}\|_2 = 1} \|\mathbf{A}\mathbf{x}\|_2$$

- Weyl's inequality: for any $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$,

$$\sigma_{k+l-1}(\mathbf{A} + \mathbf{B}) \leq \sigma_k(\mathbf{A}) + \sigma_l(\mathbf{B}), \qquad k, l \in \{1, \ldots, p\},\ k + l - 1 \leq p.$$

  Also, note the corollaries

  - $\sigma_k(\mathbf{A} + \mathbf{B}) \leq \sigma_k(\mathbf{A}) + \sigma_1(\mathbf{B})$, $k = 1, \ldots, p$
  - $|\sigma_k(\mathbf{A} + \mathbf{B}) - \sigma_k(\mathbf{A})| \leq \sigma_1(\mathbf{B})$, $k = 1, \ldots, p$
  - $\sigma_1(\mathbf{A} + \mathbf{B}) \leq \sigma_1(\mathbf{A}) + \sigma_1(\mathbf{B})$, $k = 1, \ldots, p$

Singular Value Inequalities

- (interlacing) let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B} \in \mathbb{R}^{k\times l}$ be a submatrix of $\mathbf{A}$, then $\sigma_{i+m-k+n-l}(\mathbf{A}) \le \sigma_i(\mathbf{B}) \le \sigma_i(\mathbf{A})$, $i = 1, \ldots, p-(m-k+n-l)$
  - let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B}$ be $\mathbf{A}$ with one of its rows or columns deleted, then $\sigma_{i+1}(\mathbf{A}) \le \sigma_i(\mathbf{B}) \le \sigma_i(\mathbf{A})$, $i = 1, \ldots, p-1$
  - let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B}$ be $\mathbf{A}$ with a row and a column deleted, then $\sigma_{i+2}(\mathbf{A}) \le \sigma_i(\mathbf{B}) \le \sigma_i(\mathbf{A})$, $i = 1, \ldots, p-2$

- let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $1 \le k \le p$, then

$$\sum_{i=1}^{k}\sigma_i(\mathbf{A}) = \max_{\substack{\mathbf{U}\in\mathbb{R}^{m\times k},\ \mathbf{V}\in\mathbb{R}^{n\times k} \\ \|\mathbf{u}_i\|_2=1\ \forall i,\ \mathbf{u}_i^T\mathbf{u}_j=0\ \forall i\neq j \\ \|\mathbf{v}_i\|_2=1\ \forall i,\ \mathbf{v}_i^T\mathbf{v}_j=0\ \forall i\neq j}} \sum_{i=1}^{r}\mathbf{u}_i^T\mathbf{A}\mathbf{v}_i = \max_{\substack{\mathbf{U}\in\mathbb{R}^{m\times k},\ \mathbf{V}\in\mathbb{R}^{n\times k} \\ \mathbf{U}^T\mathbf{U}=\mathbf{I} \\ \mathbf{V}^T\mathbf{V}=\mathbf{I}}} \mathrm{tr}(\mathbf{U}^T\mathbf{A}\mathbf{V})$$

- for $\mathbf{A} \in \mathbb{R}^{n\times n}$, the eigenvalues of $\mathbf{A}$ are $\lambda_i(\mathbf{A})$'s with $|\lambda_1(\mathbf{A})| \ge \ldots \ge |\lambda_n(\mathbf{A})|$ and singular values of $\mathbf{A}$ are $\sigma_i(\mathbf{A})$'s with $\sigma_1(\mathbf{A}) \ge \ldots \ge \sigma_n(\mathbf{A}) \ge 0$, then $\prod_i^k |\lambda_i(\mathbf{A})| \le \prod_i^k \sigma_i(\mathbf{A})$ for $k = 1, \ldots, n$ and the equality holds when $k = n$

- and many more...

Proof of the Eckart-Young-Mirsky Thm. by Weyl's Inequality

An application of singular value inequalities is that of proving Theorem 7.7:

- for any $\mathbf{B}$ with $\operatorname{rank}(\mathbf{B}) \leq k$, we have
  - $\sigma_l(\mathbf{B}) = 0$ for $l > k$
  - (Weyl) $\sigma_{i+k}(\mathbf{A}) \leq \sigma_i(\mathbf{A} - \mathbf{B}) + \sigma_{k+1}(\mathbf{B}) = \sigma_i(\mathbf{A} - \mathbf{B})$ for $i = 1, \ldots, p - k$
  - and consequently

$$\|\mathbf{A} - \mathbf{B}\|_F^2 = \sum_{i=1}^{p} \sigma_i(\mathbf{A} - \mathbf{B})^2 \geq \sum_{i=1}^{p-k} \sigma_i(\mathbf{A} - \mathbf{B})^2 \geq \sum_{i=k+1}^{p} \sigma_i(\mathbf{A})^2$$

- the equality above is attained if we choose $\mathbf{B} = \mathbf{A}_k$

Computation of the SVD

- assume $m \geq n$ and $\sigma_1 > \sigma_2 > \ldots \sigma_n > 0$

The power iteration can be used to compute the thin SVD, and the idea is as follows.

- form $\mathbf{A}^T\mathbf{A}$
- apply the power iteration to $\mathbf{A}^T\mathbf{A}$ to obtain $\mathbf{v}_1$
- obtain $\mathbf{u}_1 = \mathbf{A}\mathbf{v}_1/\|\mathbf{A}\mathbf{v}_1\|_2, \sigma_1 = \|\mathbf{A}\mathbf{v}_1\|_2$ (why is this true?)
- do deflation $\mathbf{A} := \mathbf{A} - \sigma_1\mathbf{u}_1\mathbf{v}_1^T$, and repeat the above steps until all singular components are found

Computation of the SVD

The QR iteration can be used to compute the thin SVD, and the idea is as follows.

- form $\mathbf{A}^T\mathbf{A}$
- apply the (symmetric) QR iteration to obtain the eigendec. $\mathbf{A}^T\mathbf{A} = \mathbf{V}_1\tilde{\mathbf{\Sigma}}^2\mathbf{V}_1^T$
- solve $\mathbf{U}\mathbf{\Sigma} = (\mathbf{A}\mathbf{V}_1)\mathbf{\Pi}$ via QR factorization with column pivoting where $\mathbf{\Sigma} \in \mathbb{R}^{m\times n}$ is a diagonal matrix with diagonal entries being the nonnegative square root of diagonal entries of $\tilde{\mathbf{\Sigma}}^2$

Remark: this approach is numerically unstable which depends on the $(\kappa(\mathbf{A}))^2$ (just as the issue in using the methods of normal equations for certain LS problems)

Computation of the SVD

- Associated with any $\mathbf{A}$ is the real symmetric matrix $\mathbf{A}^T\mathbf{A}$, whose eigenvalues tell us what the singular values of $\mathbf{A}$ are, but the relationship between the eigenvalues of $\mathbf{A}^T\mathbf{A}$ and the singular values of $\mathbf{A}$ is nonlinear.

- another real symmetric matrix assoc. with $\mathbf{A}$ has better properties in this regard

- let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and define the real symmetric matrix

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \in \mathbb{S}^{m+n}$$

  - matrix $\mathbf{J}$ is called the Jordan-Wielandt matrix
  - eigenvalues of $\mathbf{J}$ are $\pm\sigma_1(\mathbf{A}), \ldots, \pm\sigma_p(\mathbf{A})$ together with $|m-n|$ zeros
  - eigenvector of $\mathbf{J}$ associated with $\pm\sigma_i(\mathbf{A})$ $(i = 1, \ldots, p)$ is $\frac{1}{\sqrt{2}}[\ \mathbf{v}_i^T \ \ \pm\mathbf{u}_i^T\ ]^T$

Computation of the SVD

- if $m \geq n$, $\mathbf{J}$ obtains an eigendecomposition given by

$$\mathbf{J} = \mathbf{Q}\mathrm{Diag}(\sigma_1(\mathbf{A}), \ldots, \sigma_p(\mathbf{A}), -\sigma_1(\mathbf{A}), \ldots, -\sigma_p(\mathbf{A}), \underbrace{0, \ldots, 0}_{m-n \text{ zeros}})\mathbf{Q}^T$$

  where $\mathbf{Q}$ is

$$\mathbf{Q} = \frac{1}{\sqrt{2}}\begin{bmatrix} \mathbf{V} & \mathbf{V} & \mathbf{0} \\ \mathbf{U}_1 & -\mathbf{U}_1 & \sqrt{2}\mathbf{U}_2 \end{bmatrix}$$

- **Fact:** by applying symmetric QR iteration to $\mathbf{J}$ to find $\mathbf{U}$ and $\mathbf{V}$, we are *implicitly* computing the QR iteration of $\mathbf{A}^T\mathbf{A}$

- standard method to compute SVD from results for eigenvalues of real symmetric matrices

Computation of the SVD

Algorithm: SVD via Symmetric QR Iteration
input: $\mathbf{A} \in \mathbb{R}^{m \times n}$ $(m \geq n)$
form $\mathbf{J}$
$[\mathbf{Q}, \mathbf{\Lambda}] =$SymQRIteration$(\mathbf{J})$ % symmetric QR iteration
obtain $\mathbf{U}$ and $\mathbf{V}$ from $\mathbf{Q}$
obtain $\mathbf{\Sigma}$ from $\mathbf{\Lambda}$
output: $\mathbf{U}$, $\mathbf{\Sigma}$, $\mathbf{V}$

- in Topic 5, to reduce the computation cost in Hermitian eigenvalue problems
  1. apply orthogonal transformations to obtain a tridiagonal form (an upper Hessenberg form for general $\mathbf{A}$)
     (Recall: any $\mathbf{A} \in \mathbb{H}^n$ can be unitarily transformed to a tridiagonal form as $\mathbf{T} = \mathbf{V}_T^T \mathbf{A} \mathbf{V}_T$, but a diagonal form is not attainable)
  2. diagonalize the tridiagonal form by, say, the symmetric QR iteration

- since $\mathbf{J}$ is symmetric, apply tradiagonal reduction aforehead can be desirable

Computation of the SVD

- Fact: any $\mathbf{A} \in \mathbb{R}^{m \times n}$ can be unitarily transformed to an upper bidiagonal form as $\mathbf{B} = \mathbf{U}_B^T \mathbf{A} \mathbf{V}_B$ where $\mathbf{B}$ is upper bidiagonal, but a diagonal form is not attainable

- it is easy to show if $\mathbf{B}$ is bidiagonal then $\mathbf{B}^T \mathbf{B}$ is symmetric tridiagonal
  - the bidagonal reduction of $\mathbf{A}$ is related to the tridiagonal reduction of $\mathbf{A}^T \mathbf{A}$

- for $\mathbf{A} \in \mathbb{R}^{m \times n}$ $(m \geq n)$, the standard method for SVD computation is
  1. apply orthogonal transformations to abtain a upper bidiagonal form
  2. diagonalize the bidiagonal form

$$\mathbf{A} = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix} \xrightarrow{\text{Stage 1}} \begin{bmatrix} \times & \times & & \\ & \times & \times & \\ & & \times & \times \\ & & & \times \\ & & & \end{bmatrix} \xrightarrow{\text{Stage 2}} \begin{bmatrix} \times & & & \\ & \times & & \\ & & \times & \\ & & & \times \\ & & & \end{bmatrix}$$

Computation of the SVD

- **Bidiagonal reduction:** applying Householder reflectors alternately on the left and right
  - left reflector introduces zeros below the diagonal
  - right reflector introduces a row of zeros to the right of the first superdiagonal

$$\mathbf{A} = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix} \xrightarrow{\mathbf{U}_1^T} \underbrace{\begin{bmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}}_{\tilde{\mathbf{A}}_1 = \mathbf{U}_1^T \mathbf{A}} \xrightarrow{\mathbf{V}_1} \underbrace{\begin{bmatrix} \times & \times & 0 & 0 \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & \times & \times & \times \end{bmatrix}}_{\mathbf{A}_1 = \mathbf{U}_1^T \mathbf{A} \mathbf{V}_1} \to \ldots$$

  - $\mathbf{U}_1^T$ is the Householder reflector that reflects $\mathbf{A}(1:m,1)$
  - $\mathbf{V}_1 = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{V}}_1 \end{bmatrix}$ with $\tilde{\mathbf{V}}_1$ the Householder reflector that reflects $\tilde{\mathbf{A}}_1(1, 2:n)$

Computation of the SVD

- finally, we obtain

$$\underbrace{\mathbf{U}_n^T \mathbf{U}_{n-1}^T \cdots \mathbf{U}_1^T}_{\mathbf{U}_B^T} \mathbf{A} \underbrace{\mathbf{V}_1 \mathbf{V}_2 \cdots \mathbf{V}_{n-2}}_{\mathbf{V}_B} = \mathbf{B}$$

where $\mathbf{B}$ is a bidiagonal matrix that has the form

$$\mathbf{B} = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ & \alpha_2 & \ddots & \\ & & \ddots & \beta_{n-1} \\ & & & \alpha_n \\ & & & \\ & & & \\ & & & \end{bmatrix} \in \mathbb{R}^{m \times n}$$

and it can be verified that $\alpha_i \geq 0$ and $\beta_i \geq 0$

- complexity: $\mathcal{O}(4mn^2)$
- also called Golub-Kahan bidiagonalization

Computation of the SVD

- SVD of bidiagonal form $\mathbf{B}$: the task is to solve a real symmetric eigenvalue problem for $\mathbf{B}^T\mathbf{B}$, $\mathbf{B}\mathbf{B}^T$, or $\mathbf{J}_B = \begin{bmatrix} \mathbf{0} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{bmatrix}$
  - permutations are applied so that $\mathbf{\Pi} \begin{bmatrix} \mathbf{0} & \mathbf{B}^T \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \mathbf{\Pi}^T$ is symmetric tridiagonal, and then methods for symmetric tridiagonal eigenvalue problems such as divide-and-conquer (cf. Chapter 8.3-8.5 of **[Golub-Van Loan'13]**) can be used
  - implicit QR iteration for $\mathbf{B}^T\mathbf{B}$ or $\mathbf{B}\mathbf{B}^T$ by directly working on $\mathbf{B}$ (cf. Chapter 8.6.3 of **[Golub-Van Loan'13]**)

- after we get the SVD $\mathbf{B} = \tilde{\mathbf{U}}\mathbf{\Sigma}\tilde{\mathbf{V}}^T$, the SVD for $\mathbf{A}$ is given by

$$\mathbf{A} = \underbrace{\mathbf{U}_B\tilde{\mathbf{U}}}_{\mathbf{U}}\mathbf{\Sigma}\underbrace{\tilde{\mathbf{V}}^T\mathbf{V}_B^T}_{\mathbf{V}^T}$$

Computation of the SVD

Algorithm: SVD via Symmetric Tridiagonal QR Iteration
input: $\mathbf{A} \in \mathbb{R}^{m \times n}$
$\mathbf{B} = \mathbf{U}_B^T \mathbf{A} \mathbf{V}_B$ % bidiagonal reduction for $\mathbf{A}$
form $\mathbf{J}_B$
$[\mathbf{Q}, \mathbf{\Lambda}] =$SymTriQRIteration$(\mathbf{\Pi} \mathbf{J}_B \mathbf{\Pi}^T)$ % symmetric tridiagonal QR iteration
obtain $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ from $\mathbf{Q}$
obtain $\mathbf{\Sigma}$ from $\mathbf{\Lambda}$
$\mathbf{U} = \mathbf{U}_B \tilde{\mathbf{U}}$
$\mathbf{V} = \mathbf{V}_B \tilde{\mathbf{V}}$
output: $\mathbf{U}$, $\mathbf{\Sigma}$, $\mathbf{V}$

References

[Horn-Johnson'12]. R. A. Horn and C. R. Johnson, *Matrix analysis*, 2nd edition, Cambridge University Press, 2012.

[Recht-Fazel-Parrilo'10] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.

[Golub-Van Loan'13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th edition, JHU Press, 2013.