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

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

- a linear system $\mathbf{y} = \mathbf{A}\mathbf{x}$ is said to be consistent if $\mathbf{U}_2^T \mathbf{y} = \mathbf{0}$, i.e., $\mathbf{y} \in \mathcal{R}(\mathbf{A})$

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

Least Squares: Solution via SVD

- consider the least squares 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 $\mathbf{x}\in\mathbb{R}^n$,

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

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

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

Pseudo-Inverse

The pseudo-inverse (or Moore-Penrose inverse) of a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is defined as

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

From the above definition, we can show that

- let $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{A}^{\dagger}$ always exists and unique
- for least squares, $\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{y} + \boldsymbol{\eta}$ for any $\boldsymbol{\eta} \in \mathcal{R}(\mathbf{V}_2)$
- for linear system, $\mathbf{x} = \mathbf{A}^{\dagger}\mathbf{y} + \boldsymbol{\eta}$ for any $\boldsymbol{\eta} \in \mathcal{R}(\mathbf{V}_2)$ and $\mathbf{U}_2^T \mathbf{y} = \mathbf{0}$
- it can be easily shown that

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

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

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
  - another definition for Moore-Penrose inverse
- note: in general, $\mathbf{A}\mathbf{A}^\dagger \neq \mathbf{I}$ and $\mathbf{A}^\dagger\mathbf{A} \neq \mathbf{I}$; $\mathbf{A}\mathbf{B} = \mathbf{I} \Longrightarrow \mathbf{B} = \mathbf{A}^\dagger$ or $\mathbf{A} = \mathbf{B}^\dagger$

Pseudo-Inverse

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 $a \neq 0$
- $\mathrm{rank}(\mathbf{A}^\dagger) = \mathrm{rank}(\mathbf{A}) = \mathrm{rank}(\mathbf{A}^\dagger\mathbf{A}) = \mathrm{rank}(\mathbf{A}\mathbf{A}^\dagger)$
- $\mathbf{A}^\dagger = (\mathbf{A}^T\mathbf{A})^\dagger\mathbf{A}^T = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^\dagger$, $(\mathbf{A}^T)^\dagger = (\mathbf{A}\mathbf{A}^T)^\dagger\mathbf{A} = \mathbf{A}(\mathbf{A}^T\mathbf{A})^\dagger$
- $(\mathbf{A}\mathbf{A}^T)^\dagger = (\mathbf{A}^T)^\dagger\mathbf{A}^\dagger$, $(\mathbf{A}^T\mathbf{A})^\dagger = \mathbf{A}^\dagger(\mathbf{A}^T)^\dagger$
- $\mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^\dagger = (\mathbf{A}\mathbf{A}^T)^\dagger\mathbf{A}\mathbf{A}^T = \mathbf{A}\mathbf{A}^\dagger$, $\mathbf{A}^T\mathbf{A}(\mathbf{A}^T\mathbf{A})^\dagger = (\mathbf{A}^T\mathbf{A})^\dagger\mathbf{A}^T\mathbf{A} = \mathbf{A}^\dagger\mathbf{A}$
- for orthogonal $\mathbf{P}$, $\mathbf{Q}$, $(\mathbf{PAQ})^\dagger = \mathbf{Q}^T\mathbf{A}^\dagger\mathbf{P}^T$
- note: for $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$, in general (a) $(\mathbf{AB})^\dagger \neq \mathbf{B}^\dagger\mathbf{A}^\dagger$; (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 $\mathbf{A}$

Pseudo-Inverse

some properties of the pseudo-inverse:

- 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 $\mathbf{A}^\dagger = \mathbf{A}^{-1}$

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

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

- $\mathbf{P}_{\mathbf{A}^T}$ 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 have the following properties:

- $\mathcal{R}(\mathbf{A}\mathbf{A}^\dagger) = \mathcal{R}(\mathbf{A}\mathbf{A}^T) = \mathcal{R}(\mathbf{A}) = \mathcal{R}((\mathbf{A}^\dagger)^T) = \mathcal{R}(\mathbf{U}_1)$
- $\mathcal{R}(\mathbf{A}^\dagger\mathbf{A}) = \mathcal{R}(\mathbf{A}^T\mathbf{A}) = \mathcal{R}(\mathbf{A}^T) = \mathcal{R}(\mathbf{A}^\dagger) = \mathcal{R}(\mathbf{V}_1)$
- $\mathcal{N}(\mathbf{A}\mathbf{A}^\dagger) = \mathcal{N}(\mathbf{A}\mathbf{A}^T) = \mathcal{N}(\mathbf{A}^T) = \mathcal{N}(\mathbf{A}^\dagger) = \mathcal{R}(\mathbf{U}_2)$
- $\mathcal{N}(\mathbf{A}^\dagger\mathbf{A}) = \mathcal{N}(\mathbf{A}^T\mathbf{A}) = \mathcal{N}(\mathbf{A}) = \mathcal{N}((\mathbf{A}^\dagger)^T) = \mathcal{R}(\mathbf{V}_2)$

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 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 $0 \leq k \leq \text{rank}(\mathbf{A}) = p$, find a matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$ such that $\text{rank}(\mathbf{B}) \leq k$ and $\mathbf{B}$ best approximates $\mathbf{A}$

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

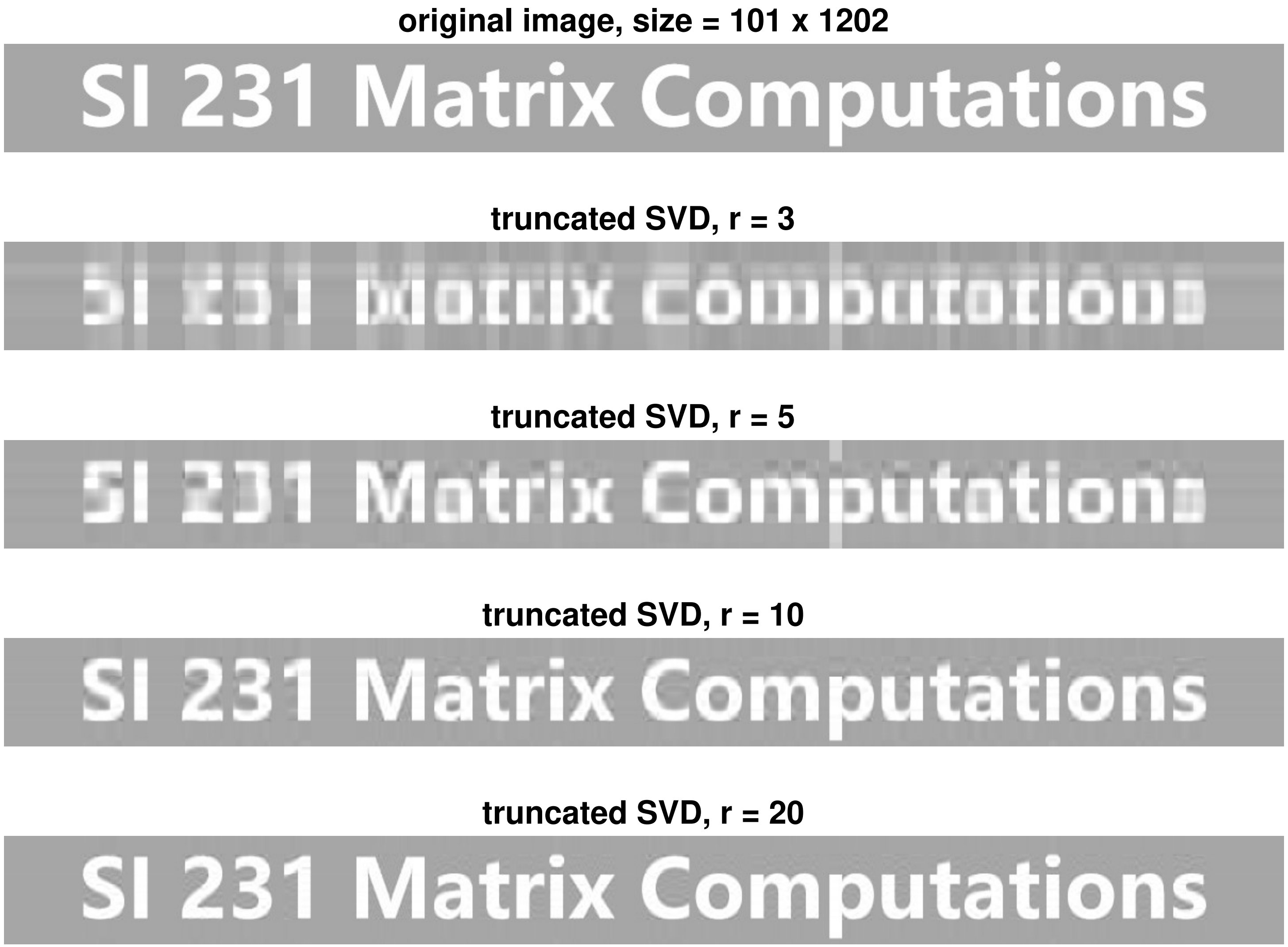
- let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be a matrix whose (i,j)th entry a_{ij} stores the (i,j)th pixel of an image

- 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 least squares sense:
  Theorem 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 by Weyl's inequality)
- also note the matrix 2-norm version of the Eckart-Young-Mirsky theorem:
  Theorem 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 the Frobenius norm or the 2-norm

Low-Rank Matrix Approximation

- recall the matrix factorization problem in Least Squares Topic:

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

  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 $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ denotes the SVD of $\mathbf{Y}$, is an optimal solution by Theorem 7

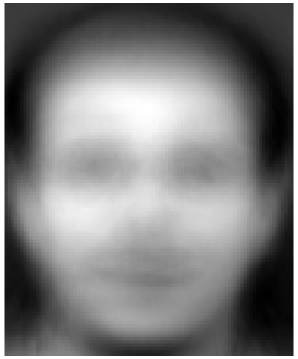
- thus, an optimal solution to the matrix factorization problem is

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

Toy Demo: Dimensionality Reduction of a Face Image Dataset

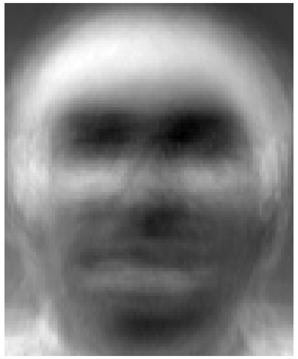
A face image dataset. Image size $= 112 \times 92$, number of face images $= 400$. Each $\mathbf{x}_i$ is the vectorization of one face image, leading to $m = 112 \times 92 = 10304$, $n = 400$.

Toy Demo: Dimensionality Reduction of a Face Image Dataset

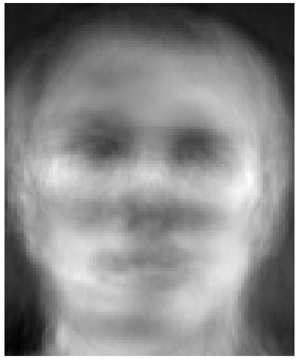


Mean face

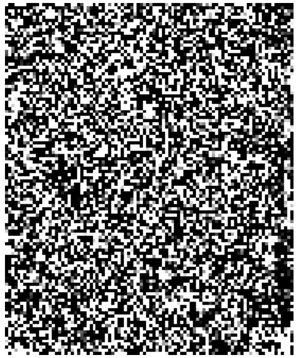
1st principal left singular vector



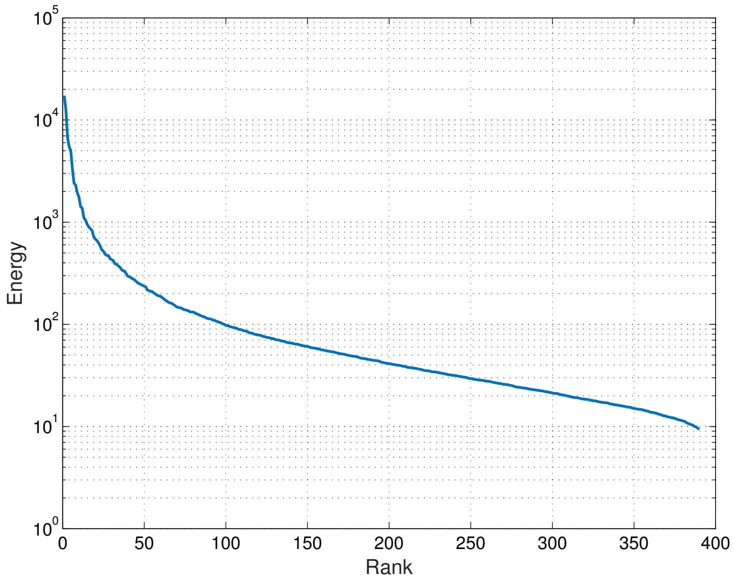
2nd principal left singular vector



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 matrices in Eigen-value Topic, we can derive various variational characterization results for singular values, e.g.,

- Courant-Fischer minimax characterization:

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

  where $\|\mathbf{A}\mathbf{x}\|_2$ can be equivalently replaced by $\|\mathbf{A}\mathbf{x}\|_2^2 = \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}$
- 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 $\sigma_{i+m-k+n-l}(\mathbf{A}) \leq \sigma_i(\mathbf{B}) \leq \sigma_i(\mathbf{A})$, $i = 1, \ldots, p - (m - k + n - l)$
  - let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B}$ be $\mathbf{A}$ with one of its rows or columns deleted, then $\sigma_{i+1}(\mathbf{A}) \leq \sigma_i(\mathbf{B}) \leq \sigma_i(\mathbf{A})$, $i = 1, \ldots, p - 1$
  - let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{B}$ be $\mathbf{A}$ with a row and a column deleted, then $\sigma_{i+2}(\mathbf{A}) \leq \sigma_i(\mathbf{B}) \leq \sigma_i(\mathbf{A})$, $i = 1, \ldots, p - 2$

- let $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $1 \leq k \leq p$, 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 $|\lambda_1(\mathbf{A})| \geq \ldots \geq |\lambda_n(\mathbf{A})|$ and singular values of $\mathbf{A}$ are $\sigma_i(\mathbf{A})$'s with $\sigma_1(\mathbf{A}) \geq \ldots \geq \sigma_n(\mathbf{A}) \geq 0$, then $\prod_i^k |\lambda_i(\mathbf{A})| \leq \prod_i^k \sigma_i(\mathbf{A})$ 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:

- for any $\mathbf{B}$ with $\text{rank}(\mathbf{B}) \leq k$, 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 $\mathbf{u}_1 = \mathbf{A}\mathbf{v}_1/\|\mathbf{A}\mathbf{v}_1\|_2, \sigma_1 = \|\mathbf{A}\mathbf{v}_1\|_2$
- 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 $\mathbf{U}_1\tilde{\mathbf{\Sigma}} = (\mathbf{A}\mathbf{V}_1)\mathbf{\Pi}$ via QR factorization with column pivoting where $\tilde{\mathbf{\Sigma}} \in \mathbb{R}^{r\times r}$ 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 least squares 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 Eigendec. Topic, to reduce the computation cost in eigenvalue problems
  1. apply orthogonal transformations to obtain a tridiagonal form for symmetric $\mathbf{A}$
     (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

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

  - $\mathbf{U}_1^T$ 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.