Matrix Computations
Chapter 6: Singular value, Singular Vectors, and Singular Value Decomposition

Section 6.1 Singular Value Decomposition

Jie Lu
ShanghaiTech University

Singular Value Decomposition

Theorem

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

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

where $\mathbf{U}$ *and* $\mathbf{V}$ *are orthogonal matrices and*

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

- 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, respectively
- Notations denoting singular values of $\mathbf{A}$:

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

Different Forms of SVD

- **Partitioned form**: Let r be the number of nonzero singular values, so that $\sigma_1 \geq \ldots \sigma_r > 0$, $\sigma_{r+1} = \ldots = \sigma_p = 0$. Then,

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

  where

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

- Thin SVD: $\mathbf{A} = \mathbf{U}_1 \tilde{\Sigma} \mathbf{V}_1^T$

- **Outer-product form**: $\mathbf{A} = \sum_{i=1}^{p} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$

SVD and Eigendecomposition

Note from the SVD $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ that

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

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

Observations:

- $(*)$ is an eigendecomposition of $\mathbf{A}\mathbf{A}^T$
- $(**)$ is an eigendecomposition of $\mathbf{A}^T\mathbf{A}$
- The left singular vector matrix $\mathbf{U}$ of $\mathbf{A}$ is the eigenvector matrix of $\mathbf{A}\mathbf{A}^T$
- The right singular vector matrix $\mathbf{V}$ of $\mathbf{A}$ is the eigenvector matrix of $\mathbf{A}^T\mathbf{A}$
- $\sigma_1^2, \ldots, \sigma_r^2$ are the nonzero eigenvalues of both $\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$

Insights of the Proof of SVD

- To see the insights of the constructive proof, consider the special case of square nonsingular $\mathbf{A}$

- $\mathbf{A}\mathbf{A}^T$ is PD with eigendecomposition

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

- Let $\Sigma = \mathrm{Diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_m})$ and $\mathbf{V} = \mathbf{A}^T\mathbf{U}\Sigma^{-1}$

- It can be verified that $\mathbf{U}\Sigma\mathbf{V}^T = \mathbf{A}$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}$

Proof of SVD

Proof of SVD (cont'd)

Proof of SVD (cont'd)

SVD and Subspaces

Properties:

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

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

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

- In practice, SVD can be used a numerical tool for computing bases of $\mathcal{R}(\mathbf{A})$, $\mathcal{R}(\mathbf{A})^\perp$, $\mathcal{R}(\mathbf{A}^T)$, $\mathcal{N}(\mathbf{A})$
- Using SVD, we can easily show the following facts:
  - $\mathrm{rank}(\mathbf{A}^T) = \mathrm{rank}(\mathbf{A})$
  - $\dim \mathcal{N}(\mathbf{A}) = n - \mathrm{rank}(\mathbf{A})$