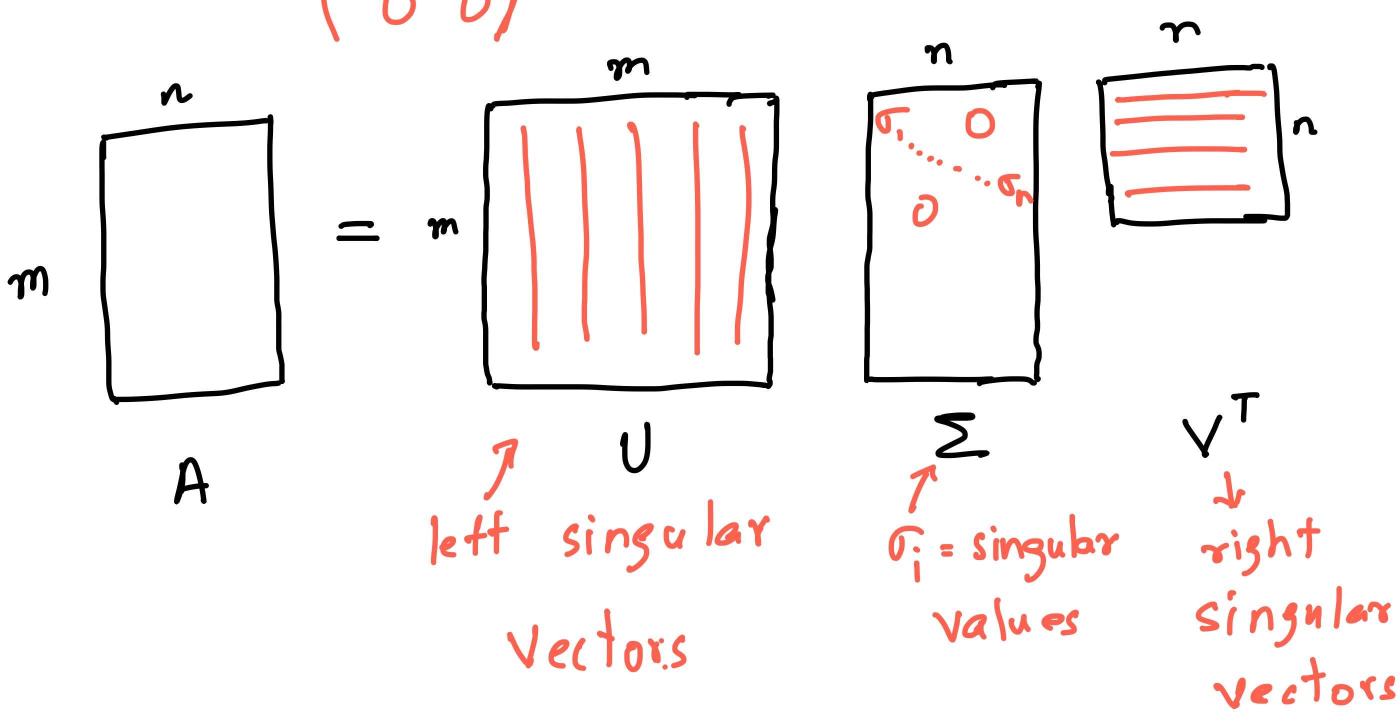
Singular Value Decomposition

Proposition: Consider $A \in \mathbb{R}^{m \times n}$ of rank r.
Then we can write A in the form

$$\boxed{A = U \cdot \Sigma \cdot V^T}$$

where $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{m \times n}$ is "diagonal" and exactly r of the diagonal values $\sigma_1, \sigma_2, \dots$ are non-zero.

$$\begin{array}{cc} & n \\ m & \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \\ 0 & & 0 \end{pmatrix} \end{array} \qquad \begin{array}{cc} & n \\ m & \begin{pmatrix} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_m & \end{pmatrix} \end{array}$$



Proof Sketch: Construct U, V, Σ, $A = U \cdot \Sigma \cdot V^T$

Given $A \in \mathbb{R}^{m \times n}$, we consider

$$B := \underbrace{A^T}_{n \times m} \underbrace{A}_{m \times n} \in \mathbb{R}^{n \times n}$$

observe :

- B is symmetric

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

- B is positive semi-definite

$$x^T B x = \langle x, Bx \rangle = \langle x, A^T A x \rangle$$
$$= \langle Ax, Ax \rangle$$
$$= \|Ax\|^2 \geq 0$$

So there exists an orthonormal basis of eigenvectors $x_1, x_2, \dots, x_n$ with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n \geq 0$.

Define:

- Σ = "$\text{diag}(\sigma_i)$" $\in \mathbb{R}^{m \times n}$

where $\sigma_i = \sqrt{\lambda_i}$

- $U = \begin{pmatrix} | \\ r_i \\ | \end{pmatrix}$ matrix with columns

$$r_i := \frac{A x_i}{\sigma_i}$$

- $V = \begin{pmatrix} | \\ x_i \\ | \end{pmatrix}$ matrix with x_i as columns.

Now we need to show that with these definitions we have $A = U \cdot \Sigma \cdot V^T$

Sketch:
- Columns of $U \cdot \Sigma$ are given as

$$\sigma_i r_i = \sigma_i \cdot \frac{A x_i}{\sigma_i} = A x_i$$

- Now multiply with V^T:
  - rows of V^T are the x_i
  - exploit that if $i \neq j$ then $x_i \perp x_j$ and $\|x_i\| = 1$
  - The terms consisting of i, j with $i \neq j$ cancel, the terms with $i = j$ will be 1.

So we will be left with the matrix A.

Key differences between SVD & eigendecomp.

- SVD always exists, no matter how A looks like.
- U, V are orthogonal (not true of eigenvectors in general)
- singular values are always real and non-negative.
- If $A \in \mathbb{R}^{n \times n}$ is symmetric, then the SVD is "nearly the same" as the eigenvalue decomposition. (λ_i, v_i) are the eigenvalue/eigenvector pairs of A, then $(|\lambda_i|, v_i)$ are the singular value/singular vector pairs of A.

In particular, left- and right singular vectors are the same.

- left-singular vectors of A are the eigenvectors of AA^T
- right- " " " " " " of $A^T A$
- $\lambda_i \neq 0$ is a eigenvalue of $A^T A$ / AA^T $\Longleftrightarrow$ $\sqrt{\lambda_i} \neq 0$ is a singular value of A.

Matrix Norms

Given a matrix $A \in \mathbb{R}^{m \times n}$. Define the following norms:

- $$\|A\|_{max} = \|A\|_\infty = \max_{ij} |a_{ij}|$$
- $$\|A\|_1 = \sum_{i,j} |a_{ij}|$$
- $$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2} = \sqrt{tr(A^T A)}$$

  Frobenius Norm

  $$= \sqrt{\sum \sigma_i^2}$$ where σ_i are the singular values of A.
- $$\|A\|_2 = \sigma_{max}(A)$$ where σ_{max} is the largest singular value.

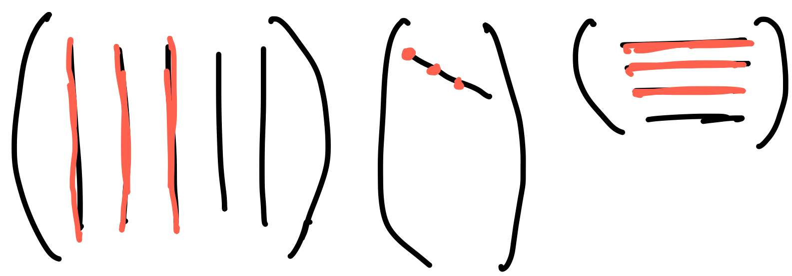
  $$= \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

  Euclidean norm on vectors in $\mathbb{R}^m$

"operator norm", "spectral norm".

Rank-K Approximation of Matrices

Given matrix $A = U \Sigma V^T$, entries $\sigma_1, \sigma_2 \ldots$ sorted in descending order. Now we are going to define a new matrix A_K as follows:

$A =$ 

- take first K cols of U

  first K entries of Σ

  first K rows of V^T

More formally:

$$A_k = \sum_{i=1}^{K} \sigma_i \underbrace{u_i v_i^T}$$

↳ rank-1 matrix

Proposition: Let B be any rank-K matrix $\in \mathbb{R}^{m \times n}$.

Then: $\|A - A_k\|_F \leq \|A - B\|_F$

"A_K is the best rank-K approximation (in Frobenius norm)"

Proposition: For any matrix B of rank-k, $B \in \mathbb{R}^{m \times n}$, $\|A - A_k\|_2 \leq \|A - B\|_2$. where $\|\cdot\|_2$ denotes the operator norm.

"A_k is the best rank-k approximation (in operator norm)

Pseudo - Inverse of Matrix

Def: For $A \in \mathbb{R}^{m \times n}$, a pseudo inverse of A is defined as the matrix $A^+ \in \mathbb{R}^{n \times m}$ which satisfies the following properties:

(i) $\underbrace{A A^+}_{\neq Id \text{ in general}} A = A$, with $\overbrace{A^+ A}^{\neq Id \text{ in general}}$

(ii) $A^+ A A^+ = A^+$

"nearly inverse" (i)–(ii)

(iii) $(A A^+)^T = A A^+$

(iv) $(A^+ A)^T = A^+ A$

symmetry (iii)–(iv)

Intuition: · A is a projection from $\mathbb{R}^3 \to \mathbb{R}^2$

$$A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

· cannot invert, obviously (inverting means reconstructing original)

· But I could "make up" a reconstruction.

$$R: \mathbb{R}^2 \to \mathbb{R}^3$$

$$R\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 5 \end{pmatrix}$$

- Now we have: $ARA = A$

$$\Downarrow$$

$$AA^+A = A$$

Proposition: Let $A \in \mathbb{R}^{m \times n}$, $A = U\Sigma V^T$ its SVD. Then: $A^+ = V\Sigma^+U^T$ where $\Sigma^+ \in \mathbb{R}^{m \times n}$.

$$\Sigma^+_{ii} = \begin{cases} 1/\Sigma_{ii} & \text{if } \Sigma_{ii} \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\Sigma = \begin{pmatrix} \sigma_1 \cdots \sigma_n \\ 0 \end{pmatrix}, \quad \Sigma^+ = \begin{pmatrix} 1/\sigma_1 & & \\ & \ddots & 1/\sigma_n \\ 0 & & \end{pmatrix}$$

Intuition: Assume $A \in \mathbb{R}^{n \times n}$, invertible, assume it has eigendecomposition $A = U\Lambda U^T$. Then:

- All entries of $\text{diag}(\Lambda)$ are $\neq 0$ (eigenvalues $\neq 0$)
- $A^{-1} = U\Lambda^{-1}U^T$ with $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

$$\Lambda^{-1} = \begin{pmatrix} 1/\lambda_1 & & 0 \\ & \ddots & \\ 0 & & 1/\lambda_n \end{pmatrix}.$$