

THE URV & SINGULAR VALUE DECOMPOSITIONS

- Orthogonal subspaces
- Orthogonal projectors, Orthogonal decomposition
- The URV decomposition
- The Singular Value Decomposition
- Properties of the SVD. Relations to eigenvalue problems

Orthogonal projectors and subspaces

Notation: Given a subspace \mathcal{X} of \mathbb{R}^m define:

$$\mathcal{X}^\perp = \{y \mid y \perp x, \forall x \in \mathcal{X}\}$$

► Let $Q = [q_1, \dots, q_r]$ an orthonormal basis of \mathcal{X}

 How would you obtain such a basis?

► Then define **orthogonal projector** $P = QQ^T$

Properties

- (a) $P^2 = P$ (b) $(I - P)^2 = I - P$
(c) $\text{Ran}(P) = \mathcal{X}$ (d) $\text{Null}(P) = \mathcal{X}^\perp$
(e) $\text{Ran}(I - P) = \text{Null}(P) = \mathcal{X}^\perp$

► Note that (b) means that $I - P$ is also a projector

Proof. (a), (b) are trivial

(c): Clearly $\text{Ran}(P) = \{x \mid x = QQ^T y, y \in \mathbb{R}^r\} \subseteq \mathcal{X}$. Any $x \in \mathcal{X}$ is of the form $x = Qy, y \in \mathbb{R}^r$. Take $Px = QQ^T(Qy) = Qy = x$. Since $x = Px$, $x \in \text{Ran}(P)$. So $\mathcal{X} \subseteq \text{Ran}(P)$. In the end $\mathcal{X} = \text{Ran}(P)$.

(d): $x \in \mathcal{X}^\perp \leftrightarrow (x, y) = 0, \forall y \in \mathcal{X} \leftrightarrow (x, Qz) = 0, \forall z \in \mathbb{R}^r \leftrightarrow (Q^T x, z) = 0, \forall z \in \mathbb{R}^r \leftrightarrow Q^T x = 0 \leftrightarrow QQ^T x = 0 \leftrightarrow Px = 0 \leftrightarrow x \in \text{Null}(P)$.

(e): Need to show inclusion both ways.

- $x \in \text{Null}(P) \leftrightarrow Px = 0 \leftrightarrow (I - P)x = x \rightarrow x \in \text{Ran}(I - P)$
 - $x \in \text{Ran}(I - P) \leftrightarrow \exists y \in \mathbb{R}^m \mid x = (I - P)y \rightarrow Px = P(I - P)y = 0 \rightarrow x \in \text{Null}(P)$ □
-

Result: Any $x \in \mathbb{R}^m$ can be written in a unique way as

$$x = x_1 + x_2, \quad x_1 \in \mathcal{X}, \quad x_2 \in \mathcal{X}^\perp$$

► Proof: Just set $x_1 = Px$, $x_2 = (I - P)x$

► Note:

$$\mathcal{X} \cap \mathcal{X}^\perp = \{0\}$$

► Therefore:

$$\mathbb{R}^m = \mathcal{X} \oplus \mathcal{X}^\perp$$

► Called the *Orthogonal Decomposition*

Orthogonal decomposition

► In other words $\mathbb{R}^m = P\mathbb{R}^m \oplus (I - P)\mathbb{R}^m$ or:

$$\mathbb{R}^m = \text{Ran}(P) \oplus \text{Ran}(I - P) \text{ or:}$$

$$\mathbb{R}^m = \text{Ran}(P) \oplus \text{Null}(P) \text{ or:}$$

$$\mathbb{R}^m = \text{Ran}(P) \oplus \text{Ran}(P)^\perp$$

► Can complete basis $\{q_1, \dots, q_r\}$ into orthonormal basis of \mathbb{R}^m , q_{r+1}, \dots, q_m

► $\{q_{r+1}, \dots, q_m\} = \text{basis of } \mathcal{X}^\perp. \rightarrow \text{dim}(\mathcal{X}^\perp) = m - r.$

Four fundamental subspaces - URV decomposition

Let $A \in \mathbb{R}^{m \times n}$ and consider $\mathbf{Ran}(A)^\perp$

$$\text{Property 1: } \mathbf{Ran}(A)^\perp = \mathbf{Null}(A^T)$$

Proof: $x \in \mathbf{Ran}(A)^\perp$ iff $(Ay, x) = 0$ for all y ; iff $(y, A^T x) = 0$ for all y ...

$$\text{Property 2: } \mathbf{Ran}(A^T) = \mathbf{Null}(A)^\perp$$

► Take $\mathcal{X} = \mathbf{Ran}(A)$ in orthogonal decomposition. ► Result:

$$\mathbb{R}^m = \mathbf{Ran}(A) \oplus \mathbf{Null}(A^T)$$

$$\mathbb{R}^n = \mathbf{Ran}(A^T) \oplus \mathbf{Null}(A)$$

4 fundamental subspaces

$$\mathbf{Ran}(A) \quad \mathbf{Null}(A^T)$$

$$\mathbf{Ran}(A^T) \quad \mathbf{Null}(A)$$

► Express the above with bases for \mathbb{R}^m :

$$\left[\underbrace{u_1, u_2, \dots, u_r}_{\text{Ran}(A)}, \underbrace{u_{r+1}, u_{r+2}, \dots, u_m}_{\text{Null}(A^T)} \right]$$

and for \mathbb{R}^n $\left[\underbrace{v_1, v_2, \dots, v_r}_{\text{Ran}(A^T)}, \underbrace{v_{r+1}, v_{r+2}, \dots, v_n}_{\text{Null}(A)} \right]$

► Observe $u_i^T A v_j = 0$ for $i > r$ or $j > r$. Therefore

$$U^T A V = R = \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}_{m \times n} \quad C \in \mathbb{R}^{r \times r} \quad \longrightarrow$$

$$A = U R V^T$$

► General class of URV decompositions


➤ Far from unique.

2 Show how you can get a decomposition in which C is lower (or upper) triangular, from the above factorization.

➤ Can select decomposition so that R is upper triangular → URV decomposition.

➤ Can select decomposition so that R is lower triangular → ULV decomposition.

➤ SVD = special case of URV where R = diagonal

3 How can you get the ULV decomposition by using only the Householder QR factorization (possibly with pivoting)? [Hint: you must use Householder twice]

The Singular Value Decomposition (SVD)

Theorem For any matrix $A \in \mathbb{R}^{m \times n}$ there exist unitary matrices $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ such that

$$A = U\Sigma V^T$$

where Σ is a diagonal matrix with entries $\sigma_{ii} \geq 0$.

$$\sigma_{11} \geq \sigma_{22} \geq \cdots \geq \sigma_{pp} \geq 0 \text{ with } p = \min(n, m)$$

► The σ_{ii} 's are the **singular values**. Notation change $\sigma_{ii} \longrightarrow \sigma_i$

Proof: Let $\sigma_1 = \|A\|_2 = \max_{x, \|x\|_2=1} \|Ax\|_2$. There exists a pair of unit vectors v_1, u_1 such that

$$Av_1 = \sigma_1 u_1$$

➤ Complete v_1 into an orthonormal basis of \mathbb{R}^n

$$V \equiv [v_1, V_2] = n \times n \text{ unitary}$$

➤ Complete u_1 into an orthonormal basis of \mathbb{R}^m

$$U \equiv [u_1, U_2] = m \times m \text{ unitary}$$

4 Define U, V as single Householder reflectors.

➤ Then, it is easy to show that

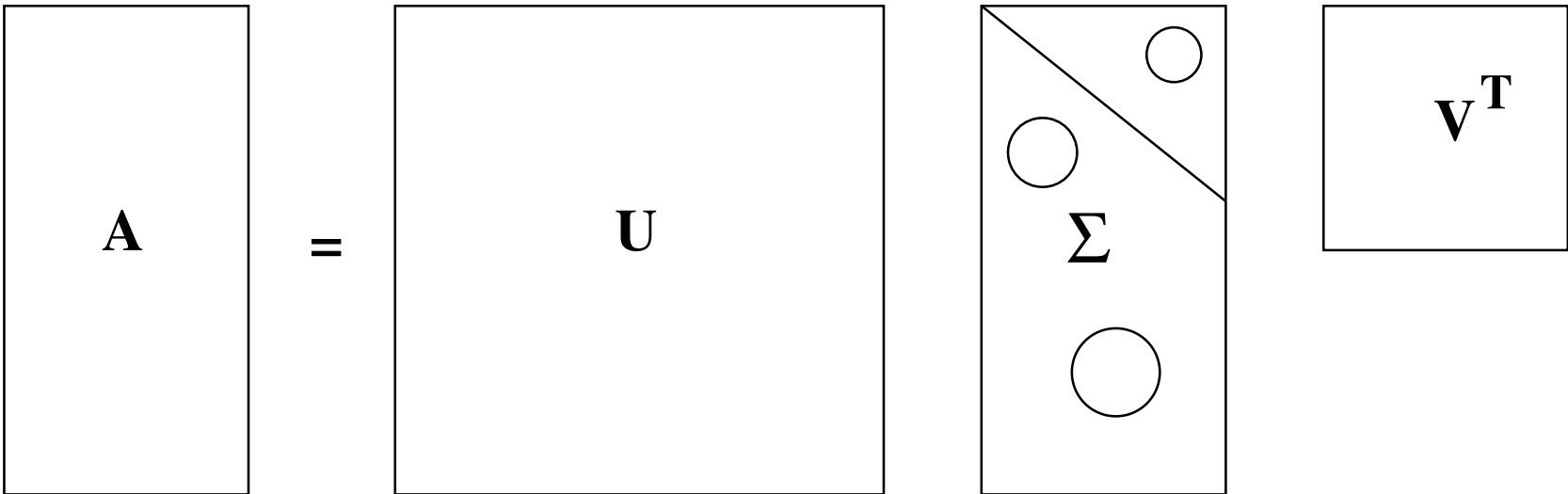
$$AV = U \times \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix} \rightarrow U^T AV = \begin{pmatrix} \sigma_1 & w^T \\ 0 & B \end{pmatrix} \equiv A_1$$

- Observe that

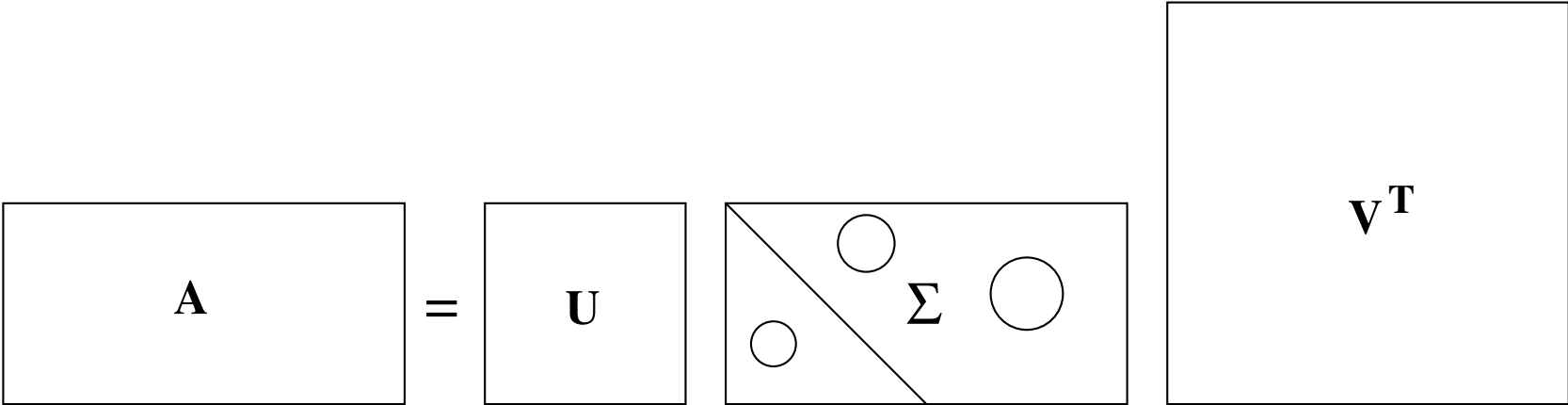
$$\left\| A_1 \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2 \geq \sigma_1^2 + \|w\|^2 = \sqrt{\sigma_1^2 + \|w\|^2} \left\| \begin{pmatrix} \sigma_1 \\ w \end{pmatrix} \right\|_2$$

- This shows that w must be zero [why?]
- Complete the proof by an induction argument. ■

Case 1:



Case 2:



The “thin” SVD

- Consider the Case-1. It can be rewritten as

$$A = [U_1 U_2] \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} V^T$$

Which gives:

$$A = U_1 \Sigma_1 V^T$$

where U_1 is $m \times n$ (same shape as A), and Σ_1 and V are $n \times n$

- Referred to as the “thin” SVD. Important in practice.

 5 How can you obtain the thin SVD from the QR factorization of A and the SVD of an $n \times n$ matrix?

A few properties.

Assume that

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \text{ and } \sigma_{r+1} = \cdots = \sigma_p = 0$$

Then:

- $\text{rank}(A) = r =$ number of nonzero singular values.
- $\text{Ran}(A) = \text{span}\{u_1, u_2, \dots, u_r\}$
- $\text{Null}(A^T) = \text{span}\{u_{r+1}, u_{r+2}, \dots, u_m\}$
- $\text{Ran}(A^T) = \text{span}\{v_1, v_2, \dots, v_r\}$
- $\text{Null}(A) = \text{span}\{v_{r+1}, v_{r+2}, \dots, v_n\}$

Properties of the SVD (continued)

- The matrix A admits the SVD expansion:

$$A = \sum_{i=1}^r \sigma_i u_i v_i^T$$

- $\|A\|_2 = \sigma_1 =$ largest singular value
- $\|A\|_F = (\sum_{i=1}^r \sigma_i^2)^{1/2}$
- When A is an $n \times n$ nonsingular matrix then $\|A^{-1}\|_2 = 1/\sigma_n$

Theorem

[Eckart-Young-Mirsky] Let $k \leq r$ and $A_k = \sum_{i=1}^k \sigma_i u_i v_i^T$ then

$$\min_{\text{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

Proof: First: $\|A - B\|_2 \geq \sigma_{k+1}$, for **any** rank- k matrix B .

Consider $\mathcal{X} = \text{span}\{v_1, v_2, \dots, v_{k+1}\}$. Note:

$$\dim(\text{Null}(B)) = n - k \rightarrow \text{Null}(B) \cap \mathcal{X} \neq \{0\}$$

[Why?]

Let $x_0 \in \text{Null}(B) \cap \mathcal{X}$, $x_0 \neq 0$. Write $x_0 = Vy$. Then

$$\|(A - B)x_0\|_2 = \|Ax_0\|_2 = \|U\Sigma V^T Vy\|_2 = \|\Sigma y\|_2$$

But $\|\Sigma y\|_2 \geq \sigma_{k+1}\|x_0\|_2$ (**Show this**). $\rightarrow \|A - B\|_2 \geq \sigma_{k+1}$

Second: take $B = A_k$. Achieves the min. \square

Right and Left Singular vectors:

- v_i 's = right singular vectors;
- u_i 's = left singular vectors.

$$Av_i = \sigma_i u_i$$
$$A^T u_j = \sigma_j v_j$$

- Consequence $A^T A v_i = \sigma_i^2 v_i$ and $AA^T u_i = \sigma_i^2 u_i$
- Right singular vectors (v_i 's) are eigenvectors of $A^T A$
- Left singular vectors (u_i 's) are eigenvectors of AA^T
- Possible to get the SVD from eigenvectors of AA^T and $A^T A$ – but: difficulties due to non-uniqueness of the SVD

Define the $r \times r$ matrix

$$\Sigma_1 = \text{diag}(\sigma_1, \dots, \sigma_r)$$

► Let $A \in \mathbb{R}^{m \times n}$ and consider $A^T A \in \mathbb{R}^{n \times n}$:

$$A^T A = V \Sigma^T \Sigma V^T \rightarrow A^T A = V \underbrace{\begin{pmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{pmatrix}}_{n \times n} V^T$$

► This gives the spectral decomposition of $A^T A$.

➤ Similarly, U gives the eigenvectors of AA^T .

$$AA^T = U \underbrace{\begin{pmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{pmatrix}}_{m \times m} U^T$$

Important:

$A^T A = VD_1V^T$ and $AA^T = UD_2U^T$ give the SVD factors U, V up to signs!