More about SVD!

☆ "A Change of Bases" viewpoint

$$A = U\Sigma V^T \in \mathbb{R}^{m\times n}$$

Pick any $\mathbf{x} \in \mathbb{R}^n$ and consider

$$\tilde{\mathbf{x}} = V^T \mathbf{x}$$

Then $\tilde{\mathbf{x}}$ is the expansion coefficient of $\mathbf{x}$ w.r.t. the ONB $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$

why? You should know this by now.

But, just in case,

$$\tilde{\mathbf{x}} = V^T\mathbf{x} \iff \mathbf{x} = V\tilde{\mathbf{x}}$$

$$= \underbrace{\tilde{x}_1 \mathbf{v}_1 + \cdots + \tilde{x}_n \mathbf{v}_n}_{\text{linear comb. of } \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}.}$$

//

Now, let $\mathbf{b} = A\mathbf{x} \in \mathbb{R}^m$

Expand $\mathbf{b}$ w.r.t. the ONB $\{\mathbf{u}_1, \cdots, \mathbf{u}_m\}$

$$\hat{\mathbf{b}} = U^T\mathbf{b} = U^T A\mathbf{x} = U^T A V \tilde{\mathbf{x}}$$

$$= \underbrace{U^T U}_{= I_m} \Sigma \underbrace{V^T V}_{= I_n} \tilde{\mathbf{x}} = \Sigma \tilde{\mathbf{x}}$$

Now, we know that Σ is diagonal!

This again shows that

"Σ represents the essence of A in a much clearer manner!"

☆ SVD vs Eigenvalue Decomposition

Let $A \in \mathbb{R}^{m \times m}$ be diagonalizable, i.e., $\exists$ the eigenvalue decomposition:

$$A = X \Lambda X^{-1}$$

where $X = [\mathbb{x}_1 \ \cdots \ \mathbb{x}_m] \in \mathbb{C}^{m \times m}$

$\Lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_m) \in \mathbb{C}^{m \times m}$

satisfying $A \mathbb{x}_j = \lambda_j \mathbb{x}_j$, $j = 1, \cdots, m$

$$\Updownarrow$$

$$A X = X \Lambda$$

Note: Even if $A \in \mathbb{R}^{m \times m}$, its eigval's & eigvec's may be complex-valued! Ex. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Note that the eigenvectors $\{\mathbb{x}_1, \cdots, \mathbb{x}_m\}$ form a basis of $\mathbb{C}^m$, but not necessarily orthonormal in general unless $A^* = A$ (unitary)

Here $A^* := (\bar{a}_{ji}) = \bar{A}^T$

conjugate transposition of $A \in \mathbb{C}^{m \times m}$

"unitarity" is a generalization of "symmetry".

With the eigenvalue decomposition,

$\mathbb{b} = A \mathbb{x}$ can be simplified as

$\tilde{\mathbb{b}} = \underset{\text{diagonal}}{\Lambda} \tilde{\mathbb{x}}$ via $\begin{cases} \tilde{\mathbb{b}} = X^{-1} \mathbb{b} \\ \tilde{\mathbb{x}} = X^{-1} \mathbb{x} \end{cases}$

change of bases again!

So, we can summarize as follows:

- SVD: Use two different ONB's U, V and work for <u>any</u> matrix.
- EIG: Use one basis (not ONB in general) and work <u>only for square matrices</u>.

☆ Matrix Properties via SVD

Let $A \in \mathbb{R}^{m \times n}$,

$p := \min(m, n)$

$r := \#$ nonzero singular values $\leq p$.

<u>Thm</u> $\operatorname{rank}(A) = r$.

(Proof) Let $A = U \Sigma V^T$.

Since U, V are orthogonal matrices, they are of full rank.

Hence, $\operatorname{rank}(A) = \operatorname{rank}(\Sigma)$

$= \#$ nonzero diagonal entries

$= r$ ///

Recall $\langle u_1, \cdots, u_r \rangle := \operatorname{span}\{u_1, \cdots, u_r\}$

<u>Thm</u> $\operatorname{range}(A) = \langle u_1, \cdots, u_r \rangle$

$\operatorname{null}(A) = \langle v_{r+1}, \cdots, v_n \rangle$

(Proof) Since $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal with only r nonzero entries,

$$\text{range}(\Sigma) = \langle \mathbf{e}_1, \cdots, \mathbf{e}_r \rangle \subset \mathbb{R}^m$$

$$\Leftrightarrow \text{range}(A) = \langle \mathbf{u}_1, \cdots, \mathbf{u}_r \rangle \subset \mathbb{R}^m. \checkmark$$

On the other hand, it is clear that for any vector $\mathbf{x} \in \mathbb{R}^n$ s.t.

$$\mathbf{x} = [\underbrace{0, 0, \cdots, 0}_{r}, x_{r+1}, \cdots, x_n]^T,$$

$$\Sigma \mathbf{x} = \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ x_{r+1} \\ \vdots \\ x_n \end{bmatrix} = \mathbf{0}.$$

So, $\text{null}(\Sigma) = \langle \mathbf{e}_{r+1}, \cdots, \mathbf{e}_n \rangle \subset \mathbb{R}^n$

Then, for such $\mathbf{x}$, we have

$$A V \mathbf{x} = U \Sigma V^T V \mathbf{x} = U \Sigma \mathbf{x} = \mathbf{0}$$

i.e., any member of null(A) should be of the form $V\mathbf{x}$, $\mathbf{x} \in \text{null}(\Sigma)$

i.e., $\text{null}(A) = \langle \mathbf{v}_{r+1}, \cdots, \mathbf{v}_n \rangle \subset \mathbb{R}^n$ //

<u>Thm</u> $\|A\|_2 = \sigma_1$, $\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$

(Proof) Since U, V are orthogonal,

$$\|A\|_2 = \|\Sigma\|_2 = \max_{1\le j\le r} \{|\sigma_j|\} = \sigma_1 \checkmark$$

The Frobenius norm is also invariant w.r.t. rotations (ortho. matrix multiplications)

Hence, $\|A\|_F = \|\Sigma\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$ ///

<u>Thm</u> The nonzero singular values of A are the square roots of the nonzero eigenvalues of A^TA or AA^T.

(Proof) $A^TA = (U\Sigma V^T)^T(U\Sigma V^T)$
$= V\Sigma^T\Sigma V^T$

$$\Leftrightarrow \quad (A^TA)V = V(\underbrace{\Sigma^T\Sigma}_{\begin{array}{c} = \mathrm{diag}(\sigma_1^2,\cdots,\sigma_r^2,0,\cdots,0) \\ \in \mathbb{R}^{n\times n}\end{array}})$$

So, the col's of V are the eigenvectors of A^TA and their nonzero eigval's are $\sigma_1^2, \cdots, \sigma_r^2$
You can show similarly that the col's of U are the eigenvec's of AA^T, and their nonzero eigval's are $\sigma_1^2, \cdots, \sigma_r^2$. ///

<u>Thm</u> $A^T = A \Rightarrow \sigma_i(A) = |\lambda_i(A)|$

(Proof) HW #3 Prob 3 says:

Any symmetric matrix has only real-valued eigenvalues and the eigenvec's form an ONB.

So, $A = Q\Lambda Q^T$, $\quad Q$: ortho, Λ: diag
$= Q|\Lambda|\,\mathrm{sgn}(\Lambda)\,Q^T$

where $$|\Lambda| := \begin{bmatrix} |\lambda_1| & & 0 \\ & \ddots & \\ 0 & & |\lambda_m| \end{bmatrix}$$

$$\mathrm{sgn}(\Lambda) := \begin{bmatrix} \mathrm{sgn}(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \mathrm{sgn}(\lambda_m) \end{bmatrix}$$

Now, it's clear that $Q^T \mathrm{sgn}(\Lambda)$ is orthogonal if Q is orthogonal.

<u>why?</u>
$$\begin{aligned} &(Q^T \mathrm{sgn}(\Lambda))(Q^T \mathrm{sgn}(\Lambda))^T \\ &= Q^T \mathrm{sgn}(\Lambda)\,\mathrm{sgn}(\Lambda)\, Q \\ &= Q^T Q = I_m \end{aligned}$$

So, $$A = \underbrace{Q}_{U}\ \underbrace{|\Lambda|}_{\Sigma}\ \underbrace{(Q^T \mathrm{sgn}(\Lambda))^T}_{V^T}$$ ///

<u>Thm</u> For $A \in \mathbb{R}^{m\times m}$,
$$|\det(A)| = \prod_{i=1}^{m} \sigma_i = \sigma_1 \cdot \sigma_2 \cdot \cdots \cdot \sigma_m$$

(Proof) We'll use the following facts.

- $\det(AB) = \det(A)\cdot\det(B)$.
- $\det(A^T) = \det(A)$
- $\det(\mathrm{diag}(a_1, \cdots, a_m)) = \prod_{i=1}^{m} a_i$
- For any Q : orthogonal, $|\det(Q)| = 1$.

<u>why?</u> $\det(Q^TQ) = \det(Q^T)\cdot\det(Q) = (\det(Q))^2$
$= \det(I) = 1$. so, $|\det(Q)| = 1$ ✓

Then, $|\det(A)| = |\det(U\Sigma V^T)| = |\det(\Sigma)|$
$= \prod \sigma_i$ ///