Lecture 19:

Solutions to in-class exercises.

Today: Singular Value Decomposition (SVD).

Discussion: the SVD of a matrix is a factorization $\underset{m\times n}{A} = U \Sigma V^T$

Diagonal-ish: $\left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \sigma_2 & & 0 \\ 0 & & \ddots \; \sigma_s & \\ \hline & 0 & & 0 \end{array}\right]$ and $\sigma_1 \geq \dots \geq \sigma_s \geq 0$

Orthogonal

It has some of the most important applications across STEM.

The point: Setting the last few σ_i's to 0 loses very little information (they tend to be small).

How does it work?

The whole idea is finding an orthonormal basis $v_1, \dots, v_n$ of $\mathbb{R}^n$ such that $Au_1, \dots, Au_n$ are orthogonal.

For instance: if A is a 2x2 rotation matrix, any orthonormal basis will do. If A is symmetric and 2x2, we can find its orthogonal diagonalization with respect to the basis v_1, v_2. Then

$$Av_1 \cdot Av_2 = \lambda\mu (v_1 \cdot v_2) = 0$$

But what if A is, say, a shear? Then it's not so clear.

Brilliant idea: $\underbrace{\underset{n\times m}{A^T}\,\underset{m\times n}{A}}_{n\times n}$ is symmetric, so it has an orthonormal diagonalization with basis $\underset{\text{eig. } \lambda_1}{v_1}, \dots, \underset{\text{eig. } \lambda_n}{v_n}$.

Then $Av_1, \dots, Av_n$ are orthogonal! Indeed, $Av_i \cdot Av_j = v_i^T A^T A v_j = v_i^T \lambda_j v_j = \lambda_j \cdot v_i^T \cdot v_j = 0$

Furthermore $\|Av_i\| = \sqrt{Av_i \cdot Av_i} = \sqrt{(Av_i)^T (Av_i)} = \sqrt{v_i^T A^T A v_i} = \sqrt{\underset{\text{real!}}{\lambda_i} v_i^T v_i} = \sqrt{\lambda_i}$.

Important observation: A and v_i are real, so $\|Av_i\|$ is real. Therefore λ_i cannot be negative!

Definition 1: Let A be an $m \times n$ matrix. Let $\lambda_1, \dots, \lambda_n \geq 0$ be the eigenvalues of $A^T A$ (with possibly repetitions)

Then the singular values of A are $\sigma_1 = \sqrt{\lambda_1}, \dots, \sigma_n = \sqrt{\lambda_n}$.

From now on, order the singular values so that $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.

Observe: if $r = \operatorname{rank}(A)$, the first r singular values will be nonzero, and the rest 0.

Set $u_1 = \frac{Av_1}{\sigma_1}, \dots, u_r = \frac{Av_1}{\sigma_r}$, and complete $u_1, \dots, u_r$ to an orthonormal basis $u_1, \dots, u_m$ of $\mathbb{R}^m$.

Then,
$$A \cdot \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | & | & & | \\ \sigma_1 u_1 & \cdots & \sigma_r u_r & 0 & \cdots & 0 \\ | & & | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{pmatrix} \cdot \left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_n & \\ \hline & 0 & & 0 \end{array}\right)$$

Therefore,
$$A = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{pmatrix} \cdot \left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_n & \\ \hline & 0 & & 0 \end{array}\right) \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}^T$$
SVD of A

Example 1: Consider $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 1 \end{pmatrix}$. Then, $A^TA = \begin{pmatrix} 6 & 5 \\ 5 & 6 \end{pmatrix}$

char poly $(A^TA) = (6-\lambda)^2 - 25 = 36 - 12\lambda + \lambda^2 - 25 = \lambda^2 - 12\lambda + 11 = (\lambda - 1)(\lambda - 11)$

Singular values: $\sigma_1 = \sqrt{11}$, $\sigma_2 = 1$. (In order!)

$E_{11} = \operatorname{Ker}\begin{pmatrix} -5 & 5 \\ 5 & -5 \end{pmatrix} = \operatorname{Span}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)$ $\downarrow$ $v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$E_1 = \operatorname{Ker}\begin{pmatrix} 5 & 5 \\ 5 & 5 \end{pmatrix} = \operatorname{Span}\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$ $\downarrow$ $v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Finally $u_1 = \frac{Av_1}{\sigma_1} = \frac{1}{\sqrt{22}}\begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}$, $u_2 = \frac{Av_2}{\sigma_2} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$

To find u_3, consider $w = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. The vectors u_1, u_2, w form a basis, and u_1, u_2 are already orthonormal. Now $w^\perp = w - (w \cdot u_1)u_1 - (w \cdot u_2)u_2$

$$= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{3}{\sqrt{22}} \cdot \frac{1}{\sqrt{22}}\begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

$$= \frac{1}{22}\begin{pmatrix} 2 \\ 2 \\ -6 \end{pmatrix} = \frac{1}{11}\begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix} \Rightarrow u_3 = \frac{1}{\sqrt{11}}\begin{pmatrix} 1 \\ 1 \\ -3 \end{pmatrix}$$

Finally, $A = \begin{pmatrix} 3/\sqrt{22} & -1/\sqrt{2} & 1/\sqrt{11} \\ 3/\sqrt{22} & 1/\sqrt{2} & 1/\sqrt{11} \\ 2/\sqrt{22} & 0 & -3/\sqrt{11} \end{pmatrix} \begin{pmatrix} \sqrt{11} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}^T$ SVD for A

[Application on Octave]

In-class exercise: find the SVD of $\begin{pmatrix} 6 & 2 \\ -7 & 6 \end{pmatrix}$.