The Singular Value Decomposition

Capstone of the class. Fundamental application of linear algebra, to data analysis (among other things).

back to rectangular matrices

SVD will let you write any matrix in the form

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T \quad \text{where}$$

$$\sigma_1 \geq \cdots \geq \sigma_r > 0 \qquad r = \operatorname{rank}(A) \quad \text{and}$$

$\{u_1, \ldots, u_r\}$ and $\{v_1, \ldots, v_r\}$ are orthonormal.

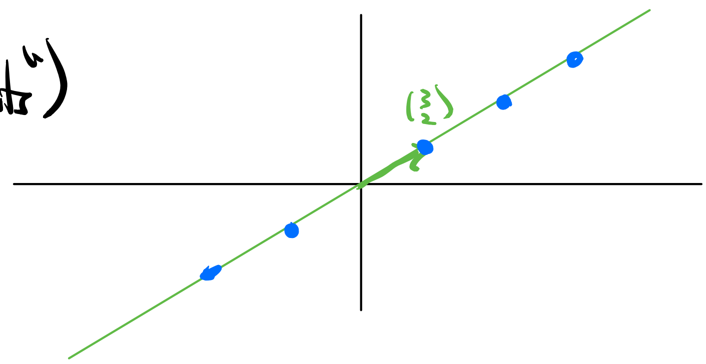
NB: Say $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$ both nonzero.

$$uv^T = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \underbrace{(v_1 \cdots v_n)}_{\text{coefficients}} = \underbrace{\begin{pmatrix} | & & | \\ v_1 u & \cdots & v_n u \\ | & & | \end{pmatrix}}_{\text{multiples of } u} = \begin{pmatrix} u_1 v_1 & \cdots & u_1 v_n \\ \vdots & \ddots & \vdots \\ u_m v_1 & \cdots & u_m v_n \end{pmatrix}$$

This is an $m \times n$ matrix of rank 1: $\operatorname{Col}(uv^T) = \operatorname{Span}\{u\}$

Let's plot the columns ("data points")

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} (-1 \quad 2 \quad 1 \quad 3 \quad -2)$$



Upshot: A rank-1 matrix encodes data points (= the columns) that lie on a line.

What about rank 2? $A = u_1 v_1^T + u_2 v_2^T$

LC's of u_1 & u_2
coeffs in v_1 & v_2

$$= \begin{pmatrix} | & & | \\ v_{11} u_1 & \cdots & v_{1n} u_1 \\ | & & | \end{pmatrix} + \begin{pmatrix} | & & | \\ v_{21} u_2 & \cdots & v_{2n} u_2 \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ v_{11} u_1 + v_{21} u_2 & \cdots & v_{1n} u_1 + v_{2n} u_2 \\ | & & | \end{pmatrix}$$

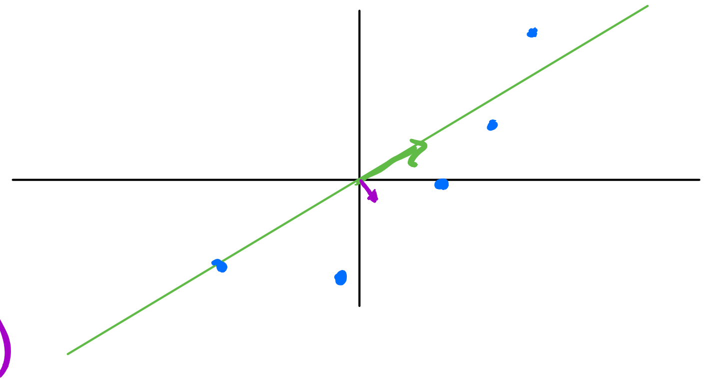
Plot the columns: coeffs of $\begin{pmatrix}3\\2\end{pmatrix}$

$$\begin{pmatrix}3\\2\end{pmatrix}\begin{pmatrix}-1 & 2 & 1 & 3 & -2\end{pmatrix}$$

orthogonal

$$+\begin{pmatrix}.2\\-.3\end{pmatrix}\begin{pmatrix}3 & 1 & 2 & -1 & 0\end{pmatrix}$$

coeffs of $\begin{pmatrix}.2\\-.3\end{pmatrix}$



Upshot: A rank-2 matrix encodes data points that lie on the plane $\mathrm{Span}\{u_1, u_2\}$

But: $\left\|\begin{pmatrix}3\\2\end{pmatrix}\right\| \gg \left\|\begin{pmatrix}.2\\-.3\end{pmatrix}\right\|$ so the $\begin{pmatrix}.2\\-.3\end{pmatrix}$ direction is less important!

$$\begin{pmatrix}3\\2\end{pmatrix}\begin{pmatrix}-1 & 2 & 1 & 3 & -2\end{pmatrix} + \begin{pmatrix}.2\\-.3\end{pmatrix}\begin{pmatrix}3 & 1 & 2 & -1 & 0\end{pmatrix}$$

$$\approx \begin{pmatrix}3\\2\end{pmatrix}\begin{pmatrix}-1 & 2 & 1 & 3 & -2\end{pmatrix} \quad \text{(to one decimal place)}$$

We've extracted important information: our data points (columns) almost lie on a line!

Applications:

- Data compression (7 numbers instead of 10 for 2×5 matrix)
- Data analysis (SVD will reveal all linear almost-relations among data points)
- Statistics (PCA: more/less important correlations)
- Quantum computing
- etc.

Outer Product Version

Thm (SVD): Let A be an $m\times n$ matrix of rank r. Then there exist $\sigma_1 \geq \cdots \geq \sigma_r > 0$ and orthonormal sets $\{u_1,\dots,u_r\}$ in $\mathbb{R}^m$ and $\{v_1,\dots,v_r\}$ in $\mathbb{R}^n$ such that

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T$$

Def:
- The σ_i are the singular values of A.
- The u_i are left singular vectors of A.
- The v_i are right singular vectors of A.

Compare: for symmetric S, we had

$$S = \lambda_1 q_1 q_1^T + \cdots + \lambda_n q_n q_n^T$$

$\lambda_1,\dots,\lambda_n$ = eigenvalues $\{q_1,\dots,q_n\}$ orthonormal eigenbasis.

NB:

$$\begin{aligned} Ax &= (\sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T)x \\ &= \sigma_1 (v_1\cdot x)u_1 + \sigma_2 (v_2\cdot x)u_2 + \cdots + \sigma_r (v_r\cdot x)u_r \end{aligned}$$

orthonormality of $\{v_1,\dots,v_r\}$

NB:

$$Av_i = \sigma_1 \overset{0}{\cancel{(v_1\cdot v_i)}}u_1 + \cdots + \sigma_i \overset{1}{\cancel{(v_i\cdot v_i)}}u_i + \cdots + \sigma_r \overset{0}{\cancel{(v_r\cdot v_i)}}u_i = \sigma_i u_i$$

So the singular vectors are related by

$\boxed{Av_i = \sigma_i u_i}$ In particular, $\boxed{\|Av_i\| = \sigma_i}$

NB: $A^T = \sigma_1 v_1 u_1^T + \cdots + \sigma_r v_r u_r^T$ is the SVD of A^T.

So A & A^T have the same:
- singular values
- singular vectors (switch right & left)

In particular, $A^T u_i = \sigma_i v_i$, so

$$A^T A v_i = A^T(\sigma_i u_i) = \sigma_i A^T u_i = \sigma_i^2 v_i$$

$$A A^T u_i = A(\sigma_i v_i) = \sigma_i A v_i = \sigma_i^2 v_i$$

This shows:

$\{v_1, \dots, v_r\}$ are orthonormal eigenvectors of $A^T A$ with eigenvalues $\sigma_1^2, \dots, \sigma_r^2$.

$\{u_1, \dots, u_r\}$ are orthonormal eigenvectors of $A A^T$ with eigenvalues $\sigma_1^2, \dots, \sigma_r^2$.

Naïve Schoolbook Procedure to Compute SVD:

Let A be an $m \times n$ matrix of rank r.

(1) Compute the nonzero eigenvalues of $A^T A$:

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$ counted w/multiplicity

(automatically r of them, & they're positive)

(2) Compute orthonormal bases of eigenspaces of $A^T A \rightsquigarrow$ orthonormal set $\{v_1, \dots, v_r\}$ s.t. $A^T A v_i = \lambda_i v_i$ for all i.

(3) Let $\sigma_i = \sqrt{\lambda_i}$ and $u_i = \frac{1}{\sigma_i} A v_i$. Then $\{u_1, \dots, u_r\}$ is orthonormal and

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T.$$

Eg: $A = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$ $\quad r = 2 \quad$ $A^TA = \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix}$

$$p(\lambda) = \lambda^2 - 50\lambda + 225 = (\lambda - 45)(\lambda - 5)$$

$$\lambda_1 = 45 \Rightarrow \sigma_1 = \sqrt{45} = 3\sqrt{5}$$

$$\lambda_2 = 5 \Rightarrow \sigma_2 = \sqrt{5}$$

$$\begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix} - 45\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \underline{-20 \quad 20} \end{pmatrix} \rightsquigarrow v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix} - 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \underline{20 \quad 20} \end{pmatrix} \rightsquigarrow v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

orthogonally diagonalize A^TA

$$u_1 = \frac{1}{3\sqrt{5}} \cdot A \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{3\sqrt{10}}\begin{pmatrix} 3 \\ 9 \end{pmatrix} = \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

$$u_2 = \frac{1}{\sqrt{5}} \cdot A \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ -1 \end{pmatrix}$$

SVD: $A = 3\sqrt{5} \cdot \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 3 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \end{pmatrix} + \sqrt{5} \cdot \frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ -1 \end{pmatrix} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \end{pmatrix}$

Check:

$$\frac{3}{2}\begin{pmatrix} 1 & 1 \\ 3 & 3 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 3 & -3 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix} = A \checkmark$$

$$\|u_1\| = \frac{1}{\sqrt{10}}\sqrt{1^2 + 3^2} = 1 \qquad \|u_2\| = \frac{1}{\sqrt{10}}\sqrt{3^2 + (-1)^2} = 1$$

$$u_1 \cdot u_2 = \frac{1}{10}\begin{pmatrix} 1 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -1 \end{pmatrix} = 0 \checkmark$$