Lecture 21

Computing SVD

Based on example from Prof. Marshall Hampton (UMN) with nice singular values

Q. What are the singular values of the matrix $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$?

Today we'll show how we compute SVD using familiar algorithms, via an example. In practice, you should just use MATLAB:

```
>> [U,S,V] = svd(A)
```

Then `A = U*S*V'`. But it's important to understand how the algorithm works; working through this example also gives an idea of

- why SVD is possible
- how it's related to previously studied decompositions
- why it's uniquely determined

Example Find the SVD of $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$ Want: $A = USV^T$

Since $A \in \mathbb{R}^{m \times n}$ where $m \le n$, we will have

$$A = U \left[\, \Sigma \;\vdots\; 0 \,\right] V^T$$

I'm actually going to compute SVD of $A^T = (USV^T)^T = VS^TU^T$ because then we work with smaller matrices.

$$A^T = V \begin{bmatrix} \Sigma \\ \hline \end{bmatrix} U^T$$

(same thing).

See the link on class website for computing SVD the other way.

Step 1: Find the singular values

Recall Singular values are the roots of the char polyn of AA^T / A^TA.

AA^T is smaller, so let's use that:

$$AA^T = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 9+4+4 & 6+6-4 \\ 6+6-4 & 4+9+4 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$ symmetric ✓

$$AA^T - \lambda I_2 = \begin{bmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{bmatrix}$$

$$\det(AA^T - \lambda I) = (17-\lambda)^2 - 64 = (x^2 - 34x + 289) - 64 = x^2 - 34 + 225$$
$$= (\lambda - 25)(\lambda - 9)$$

$$\Longrightarrow \lambda_1 = 25,\ \lambda_2 = 9 \Rightarrow \sigma_1 = 5,\ \sigma_2 = 3.$$

Step 2 Find the left singular vectors (columns of U)

(choosing to work with smaller vectors here)

- $AA^T - \lambda_1 I_2 = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix} - \begin{bmatrix} 25 & 0 \\ 0 & 25 \end{bmatrix} = \begin{bmatrix} -8 & 8 \\ 8 & -8 \end{bmatrix}$

An eigenvector of AA^T for λ_1 would be $x \in \mathbb{R}^2$ such that $(AA^T - \lambda_1 I)x = 0$. ie $x \in \ker \begin{bmatrix} -8 & 8 \\ 8 & -8 \end{bmatrix}$

We need the cols of U to be an ONB for $\mathbb{R}^2$, so choose a unit vector in $\ker \begin{bmatrix} -8 & 8 \\ 8 & 8 \end{bmatrix} = \ker \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$

Easy: $\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 - x_1 \\ x_1 - x_2 \end{bmatrix}$ This is 0 iff $x_1 = x_2$

eg. $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$$\Rightarrow \frac{x}{\|x\|} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = u_1$$

- $AA^T - \lambda_2 I = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix} - \begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix} = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}$

$\ker \begin{pmatrix} 8 & 8 \\ 8 & 8 \end{pmatrix} = \ker \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \left\langle \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\rangle \rightsquigarrow$ unit vector $\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = u_2$

Note $u_1 \perp u_2$ already! This is because AA^T is a symmetric matrix $\Rightarrow$ eigenspaces are orthogonal.

(This is because symmetric matrices are diagonalizable ie $B = Q^T \Lambda Q$ and diagonal matrices clearly have orthogonal eigenspaces...)

So the left singular vectors of A = right singular vectors of A^T are $U = [u_1 \; u_2] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$

Step 3 Since we have U already, it's easy to compute V.

Recall the outer product expansion of SVD:

$$A = USV^T = \sum \sigma_i u_i v_i^T$$

$$\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \left[\begin{array}{cc:c} 5 & 0 & 0 \\ 0 & 3 & 0 \end{array}\right] \begin{bmatrix} & V^T & \end{bmatrix}$$

$$U^T A = SV^T \Rightarrow \frac{u_i a_i^T}{\sigma_i} = v_i^T$$

$$\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \hat{V}^T$$

$$\Rightarrow \hat{V}^T = \begin{bmatrix} 1/5 & 0 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{pmatrix} 1/5 & 0 \\ 0 & 1/3 \end{pmatrix} \begin{pmatrix} 5/\sqrt{2} & 5/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 4/\sqrt{2} \end{pmatrix}$$

$$= \begin{array}{c} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/3\sqrt{2} & -1/3\sqrt{2} & 4/3\sqrt{2} \end{bmatrix} \end{array} \begin{array}{l} v_1^T \\ v_2^T \end{array}$$

For v_3, find a unit vector orthogonal to $\{v_1, v_2\}$:

$$v_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix}. \qquad v_3 \perp v_1 \Leftrightarrow a(1/\sqrt{2}) + b(1/\sqrt{2}) + 0 = 0 \Leftrightarrow a = -b.$$

$$\Rightarrow v_3 = \begin{bmatrix} a \\ -a \\ c \end{bmatrix} \qquad v_3 \perp v_2 \Leftrightarrow \underbrace{a\left(\frac{1}{\sqrt{2}}\right) - a\left(-\frac{1}{\sqrt{2}}\right)}_{2a\left(\frac{1}{\sqrt{2}}\right)} + c\left(4/\sqrt{2}\right) = 0$$

$$2a\left(\frac{1}{\sqrt{2}}\right) + 4c\left(\frac{1}{\sqrt{2}}\right) = 0$$

$$\Rightarrow 2a = -4c$$
$$c = -\frac{a}{2}.$$

$$\Rightarrow v_3 \sim \begin{bmatrix} 1 \\ -1 \\ -1/2 \end{bmatrix}. \qquad \text{Normalize: } \sqrt{1^2 + (-1)^2 + (-1/2)^2} = \sqrt{2 + \tfrac{1}{4}} = \sqrt{9/4} = 3/2.$$

$$\Rightarrow v_3 = \frac{2}{3}\begin{bmatrix} 1 \\ -1 \\ -1/2 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix} \qquad \Rightarrow V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/3\sqrt{2} & -1/3\sqrt{2} & 4/3\sqrt{2} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}$$

Conclusion:

$$\underset{A}{\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}} = \underset{U}{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}} \underset{S}{\left[\begin{array}{cc:c} 5 & 0 & 0 \\ 0 & 3 & 0 \end{array}\right]} \underset{V^T}{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/3\sqrt{2} & -1/3\sqrt{2} & 4/3\sqrt{2} \\ 2/3 & -2/3 & -1/3 \end{bmatrix}}$$

Pseudoinverse in terms of SVD

Recall we said if $A \in \mathbb{R}^{m\times n}$ is full rank, then the pseudoinverse of A is $A^+ = (A^TA)^{-1}A$.

If $A = USV^T$ $\quad \left(S = \begin{bmatrix}\Sigma \\ 0\end{bmatrix} \text{ or } \begin{bmatrix}\Sigma & 0\end{bmatrix} \right)$

then $A^+ = V S^+ U^T$

where S^+ is exactly what you'd expect:

eg. $m \geq n$

$$S = \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \\ \hline & & & 0 & & \end{bmatrix} \Rightarrow S^+ = \left[\begin{array}{cccccc|c} \frac{1}{\sigma_1} & & & & & & \\ & \ddots & & & & & \\ & & \frac{1}{\sigma_r} & & & & 0 \\ & & & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \end{array}\right]$$

This makes sense from the linear transformation perspective:

stretch

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, bend left, "V^T"] & \mathbb{R}^n \arrow[l, bend left, "V"] \arrow[r, bend left, "\tilde{\Sigma}"] & \mathbb{R}^m \arrow[l, bend left, "\tilde{\Sigma}^+"] \arrow[r, bend left, "U"] & \mathbb{R}^m \arrow[l, bend left, "U^T"] \end{tikzcd}$$

unstretch

But we can check:

① $AA^+ = (USV^T)(VS^+U^T)$

$= U \underbrace{SS^+}_{\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & 0 \\ \hline & & 0 & & \end{bmatrix}} U^T$

$= \hat{U}\hat{U}^T$

② $A^+A = (VS^+U^T)(USV^T)$

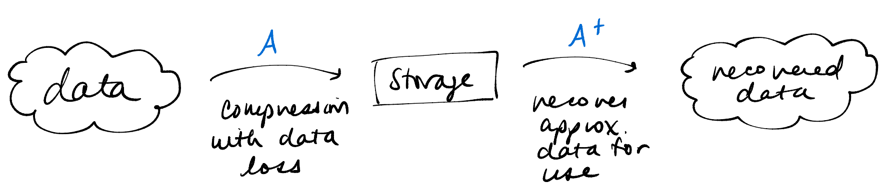
$= VS^+SV^T$

$= \hat{V}\hat{V}^T$

(from the thin decomposition)

Rmk. In many applications an inverse is not available. Pseudoinverse is the next best thing and is therefore used a lot.

cartoon idea:



Our old friend linear regression

Overdetermined system: $Ax \sim b$ where A has full column rank

$$A = USV^T = U\begin{pmatrix}\Sigma \\ 0\end{pmatrix}V^T = \left(u_1 \middle| u_2\right)\Sigma V^T$$

$U_1 = \hat{U}$

$$\|r\|^2 = \|b - Ax\|^2 = \left\|b - U\begin{pmatrix}\Sigma \\ 0\end{pmatrix}V^T x\right\|^2$$

$$= \left\|U^T\left(b - U\begin{pmatrix}\Sigma \\ 0\end{pmatrix}\underbrace{V^T x}_{y}\right)\right\|^2$$

y is a rotated/reflected x.

$$= \left\|U^T b - \begin{pmatrix}\Sigma \\ 0\end{pmatrix} y\right\|^2$$

$$= \left\|\begin{bmatrix}U_1^T b \\ U_2^T b\end{bmatrix} - \begin{bmatrix}\Sigma y \\ 0\end{bmatrix}\right\|$$

$\leftarrow$ dims in the colspace of A; we can vary y

$\leftarrow$ out of our control ($\perp$ to colspace of A)

$$\Rightarrow \|r\|^2 = \underbrace{\|U_1^T b - \Sigma y\|^2}_{\text{make this 0 to minimize } \|r\|^2} + \underbrace{\|U_2^T b\|}_{\text{nonnegotiable}}$$

If we were to have $U_1^T b - \Sigma y = 0$, we would need $U_1^T b = \Sigma y$

$\Rightarrow y = \Sigma^{-1} U_1^T b$

Σ full rank b/c A is!

ie $\forall i, \sigma_i > 0$.

$$x = Vy = V\Sigma^{-1}U_1^T b$$

$y = V^T x$

So x needs to be $V\begin{pmatrix}\sigma_1^{-1} & & & \\ & \sigma_2^{-1} & & \\ & & \ddots & \\ & & & \sigma_n^{-1}\end{pmatrix}U_1^T b$

ie $x = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i$ $\Rightarrow$ solution exists and is unique!