Singular Value Decomposition

Saturday, June 13, 2020 18:14

Compare with $S = Q \wedge Q^T$ left singular vectors

But now we do $A = U \Sigma V^T$, right singular vectors $\Sigma = \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix}$ diagonal matrix, $\sigma_i \geq 0$ called singular values.

Any matrix can be factored as $A = U \ge V^T$ The key to the math here: A^TA is nice.

Let A be $m \times n$ So, $A^{T}A = [n \times m] [m \times m] = [n \times m]$ senery ≥ 0 Symmetric matrix and PSD diagonal eigenvalues, ≥ 0 $A^{T}A = \bigvee \bigwedge \bigvee T$

VM.V

eigenvectors of ATA (orthogonal)

Courider AAT.

This is $m \times m$, but has the same non-O eigenvalues as A^TA . So, $AA^T = U \wedge U^T$

The SVD will use the U and V We're looking for $AV_1 = 0$, U_1 $AV_2 = 0$, U_2 \vdots $AV_r = 0$, U_r We could have more vectors but they would be in the null space.

: A may not be square, if may not have evectors $x \cdot s \cdot t \cdot Ax = \lambda x$ but it must have some vand $u \cdot s \cdot t \cdot Av = \sigma u$

So, we're looking for a bunch of orthogonal vectors vi s.t.

So, we're looking for a bunch of orthogonal vectors v_i s.t. Av. gives us a bunch of orthogonal vectors u_i

So,
$$A \left[v_1 \ v_2 \dots v_r \right] = \left[u_1 \ u_2 \dots u_r \right] \left[v_1 \right]$$

$$AV = U \succeq V$$

$$AVV^T = U \succeq V^T$$

$$VV^T = T$$

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

$$VV = T$$

$$VV =$$

$$A^{T}A = (U \Sigma V^{T})^{T} (U \Sigma V^{T}) = (\Sigma V^{T})^{T} U^{T}U \Sigma V^{T}$$

$$= V(\Sigma^{T} \Sigma)V^{T}$$

$$= v(\Sigma^{T} \Sigma)V^{T}$$
and σ_{i}^{2} are eigenvalues of $A^{T}A$ and also of AA^{T}

$$AA^{T} = U \ge y^{T} \forall z^{T} U^{T} = U \ge z^{T} U^{T}$$

Leigenvectors of AA^{T}

- \rightarrow Now let's see what the matrix $A = U \ge V^T$ does to a vector, as a transformation. Consider Ax
- 1. First, as VT is orthogonal, length doesn't change for x, it may just be rotated or reflected.
- 2. Consider ZVTx now: this stretches the vectors.
- 3. Vis another orthogonal matrix, so another rotation/reflection

SVD essentially tells us that every linear transformation (matrix multiplication) factors into a rotation X a stretch × a possibly different rotation.

of A is PSD and symmetric, A = Q \ Q^T

\rightarrow	to	actually	compute	the	,4 v 2	we don't	use	ATA or AAT	beoz
	it's	s very ex	pensive.	•					

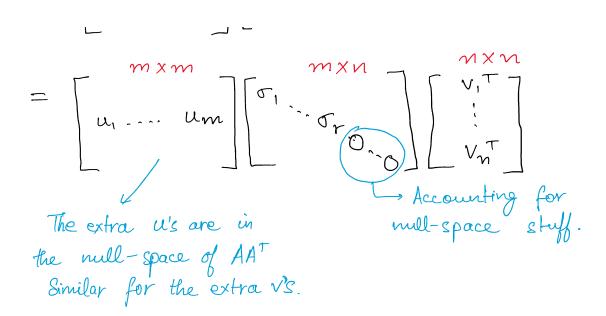
To associate with geometric interpretation:
$$\rightarrow$$
 2f A is $2x2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have 4 numbers a,b,c,d .

And we have I augle responsible for a 2D rotation + ≥ 42 numbers for 2D vector stretch 4 VT < 1 angle for 2D rotation 1+2+1=(4)

Basically for $n \times n$ (n-dimensions), we need $\binom{n}{2}$ numbers for rotation - angles b/w all pairs of axes; and n numbers for stretching. So, total: $2\binom{n}{2} + n = \frac{n^2}{L}$, no. of matrix entries!

 \Rightarrow Just like the product of eigenvalues is the determinant, for a sq. matrix $A=U \ge V^{\top}$, the product of singular values is the determinant (det $V=\det V^{\top}=1$ as V,V or thogonal).

So,
$$A = \begin{bmatrix} u_1 & \dots & u_r \end{bmatrix} \begin{bmatrix} v_1 & v_1 & v_2 & v_1 & v_2 & v_2 & v_3 & v_4 &$$



Polar decomposition

Analogous to the polar form of a complex number:

DeiO : orthogonal

S: symmetric

 \Rightarrow A = SQ (for every matrix A)

Let's see this from the SVD.

Construct a symmetric matrix: UZUT
Write

$$A = (U \ge U^{T})(UV^{T})$$

The polar decomp is used in engineering a lot.

Key fact:
If we have a big matrix of data, A and want to pull out the important parts (what data science has to do).

The biggest rank 1 part of a matrix $A = U \ge V^T$ is: $u_1 \circ_1 V_1$ Here we see A as a sum of rank I matrices:

$$A = \sum_{i=1}^{r} u_i \sigma_i v_i^{\top}$$
. Here r is the rank.