Information in the SVD

Relation to eigendecomposition

Let $Y = UDV^T$

then $YY^T = UDV^T VDU^T = UD^2U^T$

$$\left(UD^2U^T\right)u_k = UD^2\begin{pmatrix} u_1 & \text{—} \\ \vdots & \\ u_n & \text{—}\end{pmatrix}\begin{pmatrix} u_k \\ | \end{pmatrix} = UD^2\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix} \leftarrow k\text{th slot}$$

$$= U\begin{pmatrix}0\\0\\d_k^2\\0\\0\end{pmatrix} = \underset{\sim}{u}_k\, d_k^2$$

→ columns of U are eigenvectors of YY^T, w/ eigenvals $d_1^2 \dots d_n^2$

Similarly, $Y^TY = VDU^TUDV^T = VD^2V^T$

→ columns of V are eigenvectors of Y^TY, w/ eigenvalues $d_1^2 \dots d_n^2$.

Matrix rank

Def: The rank of a matrix Y is the # of lin indep columns

Rank via SVD: Let $d_1 \geq d_2 \geq \dots \geq d_r > d_{r+1} = \dots = d_n = 0$

~~(that)~~

Suppose $Y = UDV^T$, recall $V \in \mathcal{O}(n)$, $U \in \cancel{\mathcal{O}_{n,m}}$ $\mathcal{O}_m$

rotations don't change dim of row & col. space

Let $\tilde{U} = (U\ U^{\perp}) \in \mathcal{O}(m)$

$$\text{rank}(Y) = \text{rank}(\tilde{U}^TYV) = \text{rank}(\tilde{U}^TUD)$$

$$= \text{rank}\left(\begin{pmatrix} u_1 & \text{—} \\ \vdots & \\ u_n & \text{—} \\ u_1^{\perp} & \text{—} \\ \vdots & \\ u_{m-n}^{\perp} & \text{—}\end{pmatrix}\begin{pmatrix} u_1 \cdots u_n \\ | \end{pmatrix} D\right)$$

$$\text{rank}\left(\begin{pmatrix} I_{n\times n} \\ 0_{(m-n)\times n}\end{pmatrix}\begin{pmatrix} D_{r\times r} & 0_{r\times(n-r)} \\ 0_{(n-r)\times r} & 0_{(n-r)\times(n-r)}\end{pmatrix}\right) = \text{rank}\left(\begin{pmatrix} I & 0 \\ 0 & I \\ 0 & 0\end{pmatrix}\begin{pmatrix} D & 0 \\ 0 & 0\end{pmatrix}\right) = \begin{pmatrix} D_{r\times r} & 0 \\ 0 & 0\end{pmatrix}$$

r lin indep rows/co

SVD expansion:

$$Y = UDV^T = \sum_{k=1}^{r} d_k (u_k v_k^T) = \sum_{k=1}^{r} d_k M_k = \text{sum of } r \text{ "rank one matrix"}$$

Interpretation: A rank one matrix is of the form $M = uv^T$

A rank r matrix can be written as sum of r (orthogonal) rank 1 matrices

The sum of r rank one matrices has rank $\leq r$.

Sum of Squares

$$\|Y\|^2 = tr(Y^TY) = tr(VD^2V^T) = tr(D^2) = \sum d_k^2$$

$$\max_{x\neq 0} \frac{\|Yx\|}{\|x\|} = d_1 \qquad \min_{x\neq 0} \frac{\|Yx\|}{\|x\|} = d_n$$

Least Squares Matrix Approximation / Truncated SVD

Consider the following problem: $\min_{M:\, rank(M) = r} \|Y - M\|^2$

Thm: (Eckart & Young, 1936, Psychometrika)

$\|Y - M\|^2$ is minimized among rank r matrices M by $\hat{M} = \sum_{k=1}^{r} d_k u_k v_k^T$

How good is the approx?

$$Y^T\hat{M} = VDU^T U\tilde{D}V^T$$

$$\tilde{D} = \begin{pmatrix} d_1 & & & & & \\ & \ddots & & & & \\ & & d_r & & & \\ & & & 0 & & \\ & & & & 0 & \\ & & & & & 0 \end{pmatrix}$$

$$\begin{aligned}
\|Y - \hat{M}\|^2 &= tr\left((Y - \hat{M})^T(Y - \hat{M})\right) \\
&= tr(Y^TY) + tr(\hat{M}^T\hat{M}) - 2\,tr(Y^T\hat{M}) \\
&= \sum_1^n d_k^2 + \sum_1^r d_k^2 - 2\,tr(Y^T\hat{M}) \\
&= \sum_1^n d_k^2 + \sum_1^r d_k^2 - 2\,tr(D\tilde{D}) \\
&= \sum^n d_k^2 + 2\sum^r d_k^2 - 2\sum^r d_k^2 = \sum^n d^2
\end{aligned}$$

Sketch of Proof:

minimize $\|Y - uv^T\|^2 = tr\left((Y-uv^T)(Y-uv^T)^T\right)$

$= tr(YY^T) - 2tr(uv^TY^T) + tr(uv^Tvu^T)$

$= \|Y\|^2 - 2\,tr(u^TYv) + tr(u^Tu\,v^Tv)$

$= \|Y\|^2 - 2u^TYv + u^Tu\,v^Tv$

$\frac{\partial}{\partial \underset{\sim}{u}}$ " $= -2Yv + 2u\,v^Tv$

a critical point satisfies $\hat{\underset{\sim}{u}}\,(\hat{v}^T\hat{v}) = Y\hat{v}$

Sim $\hat{\underset{\sim}{v}}\,(\hat{u}^T\hat{u}) = Y^T\hat{u}$

$Y^TY\hat{v} = Y^T\hat{u}\,(\hat{v}^T\hat{v}) = \hat{v}\,(\hat{u}^T\hat{u}\cdot\hat{v}^T\hat{v}) \rightarrow \hat{v}$ is λvec of Y

$YY^T\hat{u} = \hat{u}\,(\hat{u}^T\hat{u}\cdot\hat{v}^T\hat{v}) \rightarrow \hat{u}$ is " YY