## SVD & PCA

SVD: $\underset{m\times n}{M} = \underset{m\times m}{U} \cdot \overset{\text{diagonal}}{\underset{m\times n}{\Sigma}} \cdot \underset{n\times n}{V^T}$

$U^T U = I_{m\times m} \qquad V^T V = I_{n\times n}$

Singular values: $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$

$r = \min(m,n)$

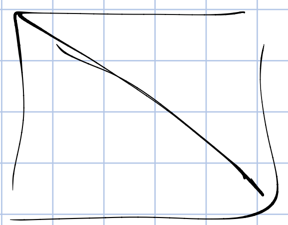
⊗ $\text{rank}(M) = \text{rank}(\Sigma) = \#$ non zero singular values

⊗ $\|M\|_F^2 = \sum_{i,j} M_{i,j}^2 = \sum_{i=1}^{r} \sigma_i^2$

Frobenius norm

$\|M\|_F^2 = \text{Tr}(M^T M) = \sum_{i=1}^{r} \lambda_i$

$\lambda_i$: eigenvalues of $M^T M$



## Low Rank Approximation

$M = U \cdot \Sigma \cdot V^T, \quad \text{rank}(M) = R$

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_R > 0$

$$\Sigma = \left(\begin{array}{ccccc|c} \sigma_1 & & & & & \\ & \sigma_2 & & 0 & & \\ & & \ddots & & & 0 \\ & & & \sigma_R & & \\ & 0 & & & 0 & \\ & & & & \ddots \; 0 & \end{array}\right)$$

$\tilde{M} = U \cdot \tilde{\Sigma} \cdot V^T \longrightarrow$ minimises $\|M - \tilde{M}\|_F$

$$\tilde{\Sigma} = \left(\begin{array}{ccccc|c} \sigma_1 & & & & & \\ & \ddots & & 0 & & \\ & & \sigma_r & & & 0 \\ & 0 & & 0 & & \\ & & & & \ddots \; 0 & \end{array}\right)$$

$\text{rank}(\tilde{M}) = r$

## CLAIM:

left singular vectors

$$U = \begin{pmatrix} \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_m \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

$$\downarrow$$

$$\hat{U}_r = \begin{pmatrix} \underline{u}_1 & \underline{u}_2 & \cdots & \underline{u}_r \\ \downarrow & \downarrow & & \downarrow \end{pmatrix} \qquad \underline{\underline{r < m}}$$

$$\tilde{M} = \underbrace{\hat{U}_r}_{m \times r} \underbrace{\hat{U}_r^T}_{r \times m} \underbrace{M}_{m \times n} \Rightarrow \text{good if } \underline{\underline{m << n}}$$

## PROOF:

$$M = \sum_{i=1}^{R} \sigma_i \underline{u}_i \underline{v}_i^T$$

$$\hat{U}_r^T \cdot M = \sum_{i=1}^{R} \sigma_i \underbrace{(\hat{U}_r^T \cdot \underline{u}_i)}_{} \cdot \underline{v}_i^T$$

$(0, \ldots, 0, 1, 0, \ldots, 0)^T$    $\underline{0}$ $i > r$

$i$    $(U^T U = I)$

$$\hat{U}_r \cdot \hat{U}_r^T \cdot M = \sum_{i=1}^{r} \sigma_i \underbrace{\left( U_r \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \right)}_{= \underline{u}_i} \cdot v_i^T = \sum_{i=1}^{r} \sigma_i \underline{u}_i \underline{v}_i^T = \tilde{M}$$

# PCA

$\underset{D\times n}{X}$ = data matrix

$D$ = dimension

$n$ = # data points.

PCA = low-dim. projection (dim $= d \ll D$)

of $X$.

= low-rank approx. for $X$.

$X = U \cdot \Sigma \cdot V^T$ (SVD)

$\hat{U}_d = \begin{pmatrix} u_1 & \text{—} & u_d \\ \downarrow & & \downarrow \end{pmatrix}$

projection: $\underset{d\times n}{Y} = \hat{U}_d^T \cdot X$

reconstruct: $\hat{X} = \hat{U}_d \hat{U}_d^T X$.