Singular value decomposition (revisited)

$$\underset{\substack{m\times d\\(m>d)}}{A} = \underset{m\times m}{V}\ \underset{m\times d}{\Sigma}\ \underset{d\times d}{U^T}$$

U and V are orthogonal, Σ is diagonal

$U = [\underline{u}_1\ \underline{u}_2\ \ \underline{u}_d]$ $\qquad$ $V = [\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_m]$

$\underline{u}_i \in \mathbb{R}^d$, orthonormal $\qquad$ $\underline{v}_i \in \mathbb{R}^m$ orthonormal

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & 0 \\ & \sigma_2 & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & 0 \end{bmatrix} \quad (r<d)$$

We calculate

$$A = [\underline{v}_1 \cdots \underline{v}_m] \begin{bmatrix} \sigma_1 & & & & 0 \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ 0 & & & & 0 \end{bmatrix} \begin{bmatrix} \underline{u}_1^T \\ \vdots \\ \underline{u}_d^T \end{bmatrix} =$$

$$= \sum_{i=1}^{r} \sigma_i \underline{v}_i \underline{u}_i^T \longleftarrow \text{Sum of rank one matrices}$$

$$= \begin{bmatrix} \sigma_1\underline{v}_1 \cdots \sigma_r\underline{v}_r \end{bmatrix} \begin{bmatrix} \underline{u}_1^T \\ \\ \underline{u}_r^T \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} \underline{v}_1 \cdots \underline{v}_r \end{bmatrix} \begin{bmatrix} \sigma_1 \underline{u}_1^T \\ \\ \sigma_r \underline{u}_r^T \end{bmatrix}$$

$$= B\,C \quad \text{where} \quad \begin{array}{l} B = m \times r \\ C = r \times d \end{array}$$

So the singular value decomposition can be interpreted as an <u>additive</u> or <u>multiplicative</u> decomposition — in both cases <u>storing</u> $(m+d)r$ <u>numbers</u> as opposed to $m \cdot d$ of the original matrix. If r is <u>much smaller</u> than d this is a <u>significant savings</u>

Suppose

$$\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r \geq \sigma_{r+1} \geq \cdots \sigma_d > 0$$

(i.e., the singular values $\sigma_{r+1} \cdots \sigma_d$ do not vanish)

Then we would instead have the approximate representations

$$A - \sum_{i=1}^{r} \sigma_i \underline{v}_i \underline{u}_i^T = \mathcal{O}(\sigma_{r+1}) \qquad \text{or}$$

$$A - \left[\sigma_1 \underline{v}_1 \cdots \sigma_r \underline{v}_r\right] \begin{bmatrix} \underline{u}_1^T \\ \\ u_r^T \end{bmatrix} = \mathcal{O}(\sigma_{r+1})$$

where $\mathcal{O}(\sigma_{r+1})$ indicates a term that is bounded by $C\sigma_{r+1}$ (say in norm) and C is a positive const (indep of σ_{r+1}).