Thus

$$\|X-A\|_F^2 \overset{(*)}{=} \|\Sigma_1\|_F^2 - 2\langle\Sigma_1, U\Sigma_2V^T\rangle_F + \|U\Sigma_2V^T\|_F^2$$

Von-Neumann's inequality.

$$\geq \|\Sigma_1\|_F^2 - 2\sum_i \sigma_i(X)\sigma_i(A)$$

$$+ \|\Sigma_2\|_F^2$$

$$= \|\Sigma_1\|_F^2 - 2\sum_i \sigma_i(X)\,\sigma_i(A) + \sum_i \sigma_i^2(A)$$

$$= \|\Sigma_1\|_F^2 + \sum_{i=1}^{d}\left[\sigma_i^2(A) - 2\sigma_i(X)\sigma_i(A)\right]$$

complete the square

$$= \|\Sigma_1\|_F^2 + \sum_{i=1}^{d}\left[(\sigma_i(A) - \sigma_i(X))^2 - \sigma_i^2(X)\right]$$

This is minimized when $\sigma_i(A) = \sigma_i(X)$, $i = 1, \cdots, d$. and

$I_{m\times m} = U = U_1^T U_2$, i.e. $U_1 = U_2$

$I_{n\times n} = V = V_2^T V_1$, i.e. $V_1 = V_2$.

Theorem/Conclusion: For X with SVD

$$X = U_1 \begin{pmatrix} \sigma_1(X) & & & & \vdots \\ & \ddots & & & \vdots \\ & & \sigma_d(X) & & \vdots \\ & & & \ddots & \vdots \\ & & & & \sigma_r(X) \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} V_1^T$$

the best rank-d approximation is

$$A = U_1 \begin{pmatrix} \sigma_1(X) & & & \vdots & \\ & \ddots & & \vdots & \\ & & \sigma_d(X) & \vdots & \\ \cdots & \cdots & \cdots & 0 & \\ & & & \vdots & 0 \end{pmatrix} V_1^T$$

Next time: midterm review.