Determining the rank of a matrix

Suppose $\boldsymbol{A}$ is a $m \times n$ rectangular matrix where $m > n$:

$$\boldsymbol{A} = \begin{pmatrix} \vdots & \dots & \vdots & \dots & \vdots \\ \boldsymbol{u}_1 & \dots & \boldsymbol{u}_n & \dots & \boldsymbol{u}_m \\ \vdots & \dots & \vdots & \dots & \vdots \end{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & 0 \\ & & \vdots \\ & & 0 \end{pmatrix} \begin{pmatrix} \dots & \mathbf{v}_1^T & \dots \\ \vdots & \vdots & \vdots \\ \dots & \mathbf{v}_n^T & \dots \end{pmatrix}$$

$$\boldsymbol{A} = \begin{pmatrix} \vdots & \dots & \vdots \\ \boldsymbol{u}_1 & \dots & \boldsymbol{u}_n \\ \vdots & \dots & \vdots \end{pmatrix} \begin{pmatrix} \dots & \sigma_1\, \mathbf{v}_1^T & \dots \\ \vdots & \vdots & \vdots \\ \dots & \sigma_n\, \mathbf{v}_n^T & \dots \end{pmatrix} = \sigma_1 \boldsymbol{u}_1 \mathbf{v}_1^T + \sigma_2 \boldsymbol{u}_2 \mathbf{v}_2^T + \cdots + \sigma_n \boldsymbol{u}_n \mathbf{v}_n^T$$

$$\boldsymbol{A} = \sum_{i=1}^{n} \sigma_i \boldsymbol{u}_i \mathbf{v}_i^T$$

$\boldsymbol{A}_1 = \sigma_1 \boldsymbol{u}_1 \mathbf{v}_1^T$ what is $\mathrm{rank}(\boldsymbol{A}_1) = ?$

In general, $\mathrm{rank}(\boldsymbol{A}_k) = k$

Rank of a matrix

For general rectangular matrix $\boldsymbol{A}$ with dimensions $m \times n$, the reduced SVD is:

$$\underbrace{\boldsymbol{A}}_{m\times n} = \underbrace{\boldsymbol{U}_{\boldsymbol{R}}}_{m\times k}\underbrace{\boldsymbol{\Sigma}_{\boldsymbol{R}}}_{k\times k}\underbrace{\boldsymbol{V}_{\boldsymbol{R}}{}^{\boldsymbol{T}}}_{k\times n}$$

$$k = \min(m, n)$$

$$\boldsymbol{A} = \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \mathbf{v}_i^T$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \\ 0 & & 0 \\ & \ddots & \vdots \\ & & 0 \end{pmatrix} \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1 & & & 0 & & \\ & \ddots & & & \ddots & \\ & & \sigma_k & 0 & \dots & 0 \end{pmatrix}$$

If $\sigma_i \neq 0 \; \forall i$, then $\text{rank}(\boldsymbol{A}) = k$ (Full rank matrix)

In general, $\text{rank}(\boldsymbol{A}) = \boldsymbol{r}$, where $\boldsymbol{r}$ is the number of non-zero singular values σ_i

$r < k$ (Rank deficient)

Rank of a matrix

- The rank of $\mathbf{A}$ equals the number of non-zero singular values which is the same as the number of non-zero diagonal elements in $\mathbf{\Sigma}$.
- Rounding errors may lead to small but non-zero singular values in a rank deficient matrix, hence the rank of a matrix determined by the number of non-zero singular values is sometimes called "effective rank".
- The right-singular vectors (columns of $\boldsymbol{V}$) corresponding to vanishing singular values span the null space of $\mathbf{A}$.
- The left-singular vectors (columns of $\boldsymbol{U}$) corresponding to the non-zero singular values of $\mathbf{A}$ span the range of $\mathbf{A}$.

SVD and Inverses

- Why is SVD so useful?
- Application #1: inverses
- $A^{-1}=(V^T)^{-1}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$
  - Using fact that inverse = transpose for orthogonal matrices
  - Since Σ is diagonal, Σ^{-1} also diagonal with reciprocals of entries of Σ

SVD and Inverses

- $A^{-1}=(V^T)^{-1}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$
- This fails when some Σ_i are 0
  - It's *supposed* to fail – singular matrix
- Pseudoinverse: if $\Sigma_i=0$, set $1/\Sigma_i$ to 0 (!)
  - "Closest" matrix to inverse
  - Defined for all (even non-square, singular, etc.) matrices
  - Equal to $(A^TA)^{-1}A^T$ if A^TA invertible

Pseudo-inverse

- **Problem:** if $\mathbf{A}$ is rank-deficient, $\mathbf{\Sigma}$ is not be invertible

How to fix it: Define the Pseudo Inverse

Pseudo-Inverse of a diagonal matrix:

$$(\mathbf{\Sigma}^{+})_i = \begin{cases} \frac{1}{\sigma_i}, & \text{if } \sigma_i \neq 0 \\ 0, & \text{if } \sigma_i = 0 \end{cases}$$

Pseudo-Inverse of a matrix $\boldsymbol{A}$:

$$\boldsymbol{A}^{+} = \boldsymbol{V}\mathbf{\Sigma}^{+}\boldsymbol{U}^{T}$$

Pseudo-inverse

If a matrix A has the singular value decomposition

$$A=UWV^T$$

then the pseudo-inverse or Moore-Penrose inverse of A is

$$A^+=V^TW^{-1}U$$

If A is 'tall' (m>n) and has full rank

$$A^+=(A^TA)^{-1}A^T$$ (it gives the least-squares solution $x_{lsq}=A^+b$)

If A is 'short' (m<n) and has full rank

$$A^+=A^T(AA^T)^{-1}$$ (it gives the least-norm solution $x_{l\text{-}n}=A^+b$)

In general, $x_{pinv}=A^+b$ is the minimum-norm, least-squares solution.

SVD and Least Squares

- Solving **Ax**=**b** by least squares
- **x**=pseudoinverse(**A**) times **b**
- Compute pseudoinverse using SVD
  - Lets you see if data is singular
  - Even if not singular, ratio of max to min singular values tells you how stable the solution will be
  - Set $1/\sum_i$ to 0 if $\sum_i$ is small (even if not exactly 0)

SVD and Least Squares

- If $\boldsymbol{A}$ is a $n \times n$ square matrix and we want to solve $\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{b}$, we can use the SVD for $\boldsymbol{A}$ such that

$$\boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T\boldsymbol{x} = \boldsymbol{b}$$
$$\boldsymbol{\Sigma}\,\boldsymbol{V}^T\boldsymbol{x} = \boldsymbol{U}^T\boldsymbol{b}$$

Solve: $\boldsymbol{\Sigma}\,\boldsymbol{y} = \boldsymbol{U}^T\boldsymbol{b}$ (diagonal matrix, easy to solve!)
Evaluate: $\boldsymbol{x} = \boldsymbol{V}\,\boldsymbol{y}$

Cost of solve: $O(n^2)$
Cost of decomposition $O(n^3)$ (recall that SVD and LU have the same cost asymptotic behavior, however the number of operations - constant factor before n^3 - for the SVD is larger than LU)

SVD and Eigenvectors

- Let A=U $\sum$V^T, and let x_i be i^{th} column of V
- Consider A^{T}A x_i:

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}x_i = \mathbf{V}\sum{}^{\mathrm{T}}\mathbf{U}^{\mathrm{T}}\mathbf{U}\sum\mathbf{V}^{\mathrm{T}}x_i = \mathbf{V}\sum{}^{2}\mathbf{V}^{\mathrm{T}}x_i = \mathbf{V}\sum{}^{2}\begin{pmatrix}0\\ \vdots\\ 1\\ \vdots\\ 0\end{pmatrix} = \mathbf{V}\begin{pmatrix}0\\ \vdots\\ \sum_i{}^2\\ \vdots\\ 0\end{pmatrix}$$

$$= \sum_i{}^2 x_i$$

- So elements of $\sum$ are sqrt(eigenvalues) and columns of V are eigenvectors of A^{T}A
  - What we wanted for robust least squares fitting!

SVD and Matrix Similarity

- One common definition for the norm of a matrix is the Frobenius norm:

$$\|\mathbf{A}\|_{\mathrm{F}} = \sum_i \sum_j a_{ij}^{\ 2}$$

- Frobenius norm can be computed from SVD

$$\|\mathbf{A}\|_{\mathrm{F}} = \sum_i \Sigma_i^{\ 2}$$

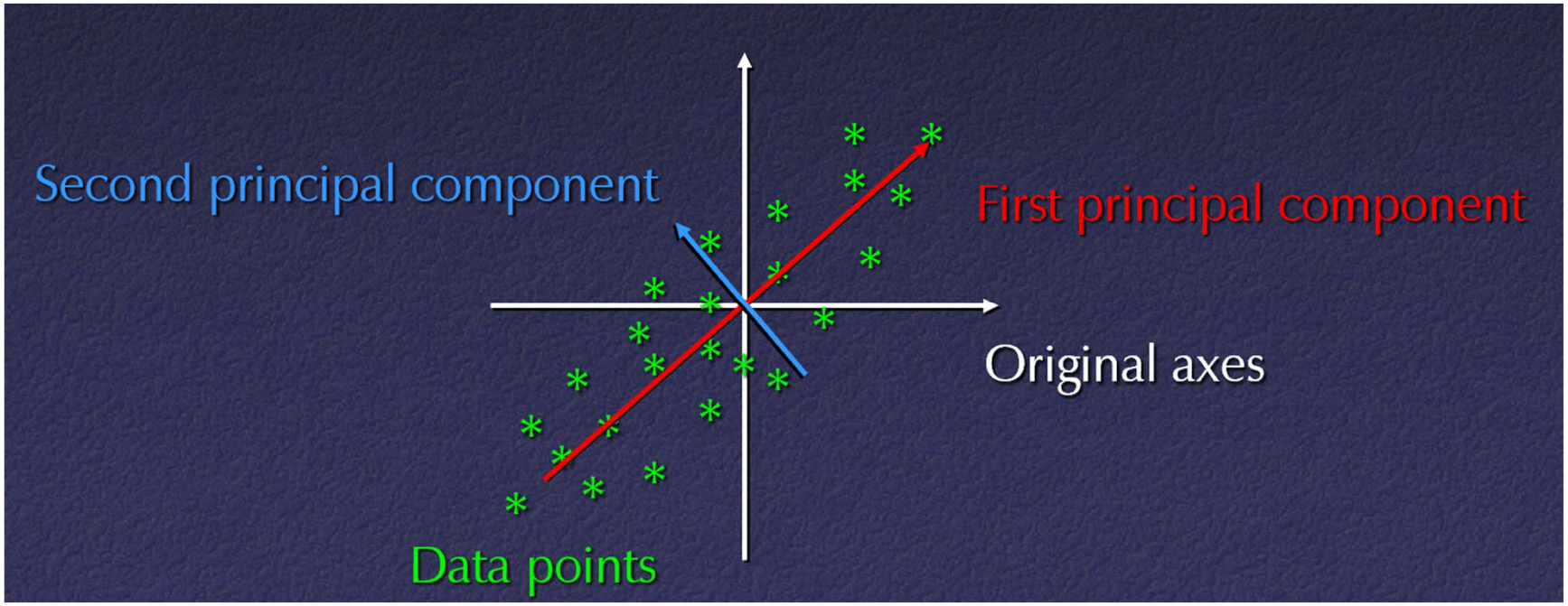
- So changes to a matrix can be evaluated by looking at changes to singular values

SVD and Matrix Similarity

- Suppose you want to find best rank-*k* approximation to A
  - Answer: set all but the largest *k* singular values to zero
- Can form compact representation by eliminating columns of U and V corresponding to zeroed $\sum_i$

SVD and PCA

- Principal Components Analysis (PCA): approximating a high-dimensional data set with a lower-dimensional subspace



SVD and PCA

- Data matrix with points as rows, take SVD
  - Subtract out mean ("whitening")
- Columns of $\mathbf{V}_k$ are principal components
- Value of $\sum_i$ gives importance of each component

Low-Rank Approximation

Another way to write the SVD (assuming for now $m > n$ for simplicity)

$$\boldsymbol{A} = \begin{pmatrix} \vdots & \cdots & \vdots \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_m \\ \vdots & \cdots & \vdots \end{pmatrix} \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_n & \\ & & & 0 \\ & & & \vdots \\ & & & 0 \end{pmatrix} \begin{pmatrix} \cdots & \mathbf{v}_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \mathbf{v}_n^T & \cdots \end{pmatrix}$$

$$= \begin{pmatrix} \vdots & \cdots & \vdots \\ \boldsymbol{u}_1 & \cdots & \boldsymbol{u}_n \\ \vdots & \cdots & \vdots \end{pmatrix} \begin{pmatrix} \cdots & \sigma_1\, \mathbf{v}_1^T & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \sigma_n\, \mathbf{v}_n^T & \cdots \end{pmatrix}$$

$$= \sigma_1 \boldsymbol{u}_1 \mathbf{v}_1^T + \sigma_2 \boldsymbol{u}_2 \mathbf{v}_2^T + \cdots + \sigma_n \boldsymbol{u}_n \mathbf{v}_n^T$$

The SVD writes the matrix A as a sum of outer products (of left and right singular vectors).

Low-Rank Approximation

The best **rank-$\boldsymbol{k}$** approximation for a $m \times n$ matrix $\boldsymbol{A}$, (where $k \leq min(m, n)$) is the one that minimizes the following problem:

$$\min_{A_k} \|A - A_k\|$$
$$\text{such that} \quad \text{rank}(A_k) \leq k.$$

When using the induced 2-norm, the best **rank-$\boldsymbol{k}$** approximation is given by:

$$\boldsymbol{A}_k = \sigma_1 \boldsymbol{u}_1 \mathbf{v}_1^T + \sigma_2 \boldsymbol{u}_2 \mathbf{v}_2^T + \cdots + \sigma_k \boldsymbol{u}_k \mathbf{v}_k^T$$

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots \geq 0$$

Note that $rank(\boldsymbol{A}) = n$ and $rank(\boldsymbol{A}_k) = k$ and the norm of the difference between the matrix and its approximation is

$$\|\boldsymbol{A} - \boldsymbol{A}_k\|_2 = \left\|\sigma_{k+1} \boldsymbol{u}_{k+1} \mathbf{v}_{k+1}^T + \sigma_{k+2} \boldsymbol{u}_{k+2} \mathbf{v}_{k+2}^T + \cdots + \sigma_n \boldsymbol{u}_n \mathbf{v}_n^T\right\|_2 = \sigma_{k+1}$$

Matrix norms

The Euclidean norm of an orthogonal matrix is equal to 1

$$\|\boldsymbol{U}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{U}\boldsymbol{x}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \sqrt{(\boldsymbol{U}\boldsymbol{x})^T(\boldsymbol{U}\boldsymbol{x})} = \max_{\|\boldsymbol{x}\|_2=1} \sqrt{\boldsymbol{x}^T\boldsymbol{x}} = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{x}\|_2 = 1$$

The Euclidean norm of a matrix is given by the largest singular value

$$\|\boldsymbol{A}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{A}\boldsymbol{x}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \left\|\boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^T\boldsymbol{x}\right\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \left\|\boldsymbol{\Sigma}\,\boldsymbol{V}^T\boldsymbol{x}\right\|_2 =$$

$$= \max_{\|\boldsymbol{V}^T\boldsymbol{x}\|_2=1} \left\|\boldsymbol{\Sigma}\,\boldsymbol{V}^T\boldsymbol{x}\right\|_2 = \max_{\|\boldsymbol{y}\|_2=1} \|\boldsymbol{\Sigma}\,\boldsymbol{y}\|_2 = \max(\sigma_i)$$

Where we used the fact that $\|\boldsymbol{U}\|_2 = 1$, $\|\boldsymbol{V}\|_2 = 1$ and $\boldsymbol{\Sigma}$ is diagonal

$$\|\boldsymbol{A}\|_2 = \max(\sigma_i) = \sigma_{max}$$

σ_{max} is the largest singular value

Norm for the inverse of a matrix

The Euclidean norm of the inverse of a square-matrix is given by:

Assume here $\boldsymbol{A}$ is full rank, so that $\boldsymbol{A}^{-1}$ exists

$$\|\boldsymbol{A}^{-1}\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \left\|(\boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{T}})^{-1}\boldsymbol{x}\right\|_2 = \max_{\|\boldsymbol{x}\|_2=1} \left\|\boldsymbol{V}\,\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{U}^{\boldsymbol{T}}\boldsymbol{x}\right\|_2$$

Since $\|\boldsymbol{U}\|_2 = 1$, $\|\boldsymbol{V}\|_2 = 1$ and $\boldsymbol{\Sigma}$ is diagonal then

$$\|\boldsymbol{A}^{-1}\|_2 = \frac{1}{\sigma_{min}}$$

σ_{min} is the smallest singular value

Norm of the pseudo-inverse matrix

- **The norm of the pseudo-inverse of a $m \times n$ matrix is:**

$$\|\boldsymbol{A}^{+}\|_2 = \frac{1}{\sigma_r}$$

where σ_r is the smallest **non-zero** singular value. This is valid for any matrix, regardless of the shape or rank.

Note that for a full rank square matrix, $\|\boldsymbol{A}^{+}\|_2$ is the same as $\|\boldsymbol{A}^{-1}\|_2$.

Zero matrix: If $\boldsymbol{A}$ is a zero matrix, then $\boldsymbol{A}^{+}$ is also the zero matrix, and $\|\boldsymbol{A}^{+}\|_2 = 0$

SVD Summary

- The SVD is a factorization of a $m \times n$ matrix into $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{\Sigma}\,\boldsymbol{V}^{\boldsymbol{T}}$ where $\boldsymbol{U}$ is a $m \times m$ orthogonal matrix, $\boldsymbol{V}^{\boldsymbol{T}}$ is a $n \times n$ orthogonal matrix and $\boldsymbol{\Sigma}$ is a $m \times n$ diagonal matrix.

- In reduced form: $\boldsymbol{A} = \boldsymbol{U_R}\boldsymbol{\Sigma_R}\boldsymbol{V_R}^{\boldsymbol{T}}$, where $\boldsymbol{U_R}$ is a $m \times k$ matrix, $\boldsymbol{\Sigma_R}$ is a $k \times k$ matrix, and $\boldsymbol{V_R}$ is a $n \times k$ matrix, and $k = \min(m, n)$.

- The columns of $\boldsymbol{V}$ are the eigenvectors of the matrix $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$, denoted the right singular vectors.

- The columns of $\boldsymbol{U}$ are the eigenvectors of the matrix $\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}$, denoted the left singular vectors.

- The diagonal entries of $\boldsymbol{\Sigma}^{\boldsymbol{2}}$ are the eigenvalues of $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$. $\sigma_i = \sqrt{\lambda_i}$ are called the singular values.

- The singular values are always non-negative (since $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$ is a positive semi-definite matrix, the eigenvalues are always $\lambda \geq 0$)

Matrix Norm

- Example
- Solution

$$\mathbf{C} = \begin{bmatrix} 3 & -2 \\ -1 & 3 \end{bmatrix}$$

$$\|\mathbf{C}\|_1 = \max_{\|\mathbf{x}\|_1=1} \|\mathbf{C}\mathbf{x}\|_1$$

$$\|\mathbf{C}\|_1 = \max_{1\le j\le 3} \sum_{i=1}^{3} |C_{ij}|$$

$$\|\mathbf{C}\|_1 = \max(|3| + |-1|, |-2| + |3|)$$

$$\|\mathbf{C}\|_1 = \max(4, 5)$$

$$\|\mathbf{C}\|_1 = 5$$

$$\|\mathbf{C}\|_\infty = \max(|3| + |-2|, |-1| + |3|)$$

$$\|\mathbf{C}\|_\infty = \max(5, 4)$$

$$\|\mathbf{C}\|_\infty = 5$$

$$\|\mathbf{C}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|\mathbf{C}\mathbf{x}\|_2$$

$$det(\mathbf{C}^T\mathbf{C} - \lambda\mathbf{I}) = 0$$

$$det(\begin{bmatrix} 3 & -1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 3 \end{bmatrix} - \lambda\mathbf{I}) = 0$$

$$det(\begin{bmatrix} 9+1 & -6-3 \\ -3-6 & 4+9 \end{bmatrix} - \lambda\mathbf{I}) = 0$$

$$det(\begin{bmatrix} 10-\lambda & -9 \\ -9 & 13-\lambda \end{bmatrix}) = 0$$

$$(10-\lambda)(13-\lambda) - 81 = 0$$

$$\lambda^2 - 23\lambda + 130 - 81 = 0$$

$$\lambda^2 - 23\lambda + 49 = 0$$

$$(\lambda - \frac{1}{2}(23 + 3\sqrt{37}))(\lambda - \frac{1}{2}(23 - 3\sqrt{37})) = 0$$

$$\|\mathbf{C}\|_2 = \sqrt{\lambda_{max}} = \sqrt{\frac{1}{2}(23 + 3\sqrt{37})} \approx 4.54$$

Matrix Norm

- Example

Show that $\|A\|_{\mathrm{F}} = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$, where $\sigma_1, \ldots, \sigma_r$ are the singular values of A. Then show that $\sigma_{\max}(A) \le \|A\|_{\mathrm{F}} \le \sqrt{r}\,\sigma_{\max}(A)$.

- Solution

$\|U\Sigma V^T\|_F$

We start with the full SVD of $A = U\Sigma V^{\mathsf{T}}$. By the previous problem,

$$\|A\|_{\mathrm{F}} = \|U\Sigma V^{\mathsf{T}}\|_{\mathrm{F}} = \|\Sigma V^{\mathsf{T}}\|_{\mathrm{F}} = \|\Sigma\|_{\mathrm{F}} = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}.$$

Since $\sigma_2^2, \ldots, \sigma_r^2 \ge 0$, we have $\sigma_1 \le \sqrt{\sigma_1^2 + \cdots + \sigma_r^2} = \|A\|_{\mathrm{F}}$. Next, since $\sigma_2, \ldots, \sigma_r \le \sigma_1$, we have $\|A\|_{\mathrm{F}} = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2} \le \sqrt{\sigma_1^2 + \cdots + \sigma_1^2} = \sqrt{r}\sigma_1$.

$(\Sigma V^T)^T(\Sigma V^T)$

!? $\|U\Sigma V^T\|_F^2 = tr(U\Sigma V^T)^T(U\Sigma V^T) = tr(V\Sigma^T \underbrace{U^T U}_{I} \Sigma V^T) = tr(V\Sigma^T\Sigma V^T) = \|\Sigma V^T\|_F^2$

چرا؟ $\|\Sigma V^T\|_F^2 = tr((\Sigma V^T)^T(\Sigma V^T)) = tr((\Sigma V^T)(V\Sigma^T))$

$= tr(\Sigma \underbrace{V^T V}_{I} \Sigma^T)$

$= tr(\Sigma\Sigma^T) = tr(\Sigma^T\Sigma)$

$= \|\Sigma\|_F$

از این خاصیت استفاده شد

$tr(AB) = tr(BA)$

Low-Rank Approximation

- Example $\mathbf{X} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$, the best rank-1 approximation is

- Solution

$$\mathbf{X}_1 = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T = \sqrt{3} \begin{pmatrix} \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

In this problem, the approximation error under either norm (spectral or Frobenius) is the same: $\|\mathbf{X} - \mathbf{X}_1\| = \sigma_2 = 1$.

Matrix norm and low-rank approximation

Let $\mathbf{A}$ be a square, symmetric matrix, with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Show that

$$\|\mathbf{A}\|_2 = \max(|\lambda_1|, |\lambda_n|)$$

Proof. It suffices to show that the singular values of $\mathbf{A}$ are given by $|\lambda_1|, \ldots, |\lambda_n|$. To see this, consider the orthogonal diagonalization of $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$. It follows that

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^2 = \mathbf{Q}\mathbf{\Lambda}^2\mathbf{Q}^T$$

Therefore, the squared singular values of $\mathbf{A}$ (which are eigenvalues of $\mathbf{A}\mathbf{A}^T$) coincide with the eigenvalues of $\mathbf{A}^2$, i.e., $\sigma_1^2 = \lambda_1^2, \ldots, \sigma_n^2 = \lambda_n^2$ (not entirely sorted). This gives that $\sigma_1 = |\lambda_1|, \ldots, \sigma_n = |\lambda_n|$ (still not sorted, but the largest singular value must be the larger one of $|\lambda_1|, |\lambda_n|$.