Matrix Decomposition

- ✓ LU
- ✓ QR
- ✓ Spectral (square matrices)
- → SVD

SVD is an extension of eigensystems to singular and rectangular matrices

Eigenproblems require that $\underline{A}$ be square and defective eigenvalues cause issues for eigen decomposition

Instead, look for the singular values σ and the vectors $\underline{u}$ and $\underline{v}$, such that

$$\underset{m\times n}{\underline{A}}\;\underset{n\times 1}{\underline{v}} = \underset{1\times 1}{\sigma}\;\underset{m\times 1}{\underline{u}}\;, \qquad \underline{A} \in \mathbb{R}^{m\times n}$$

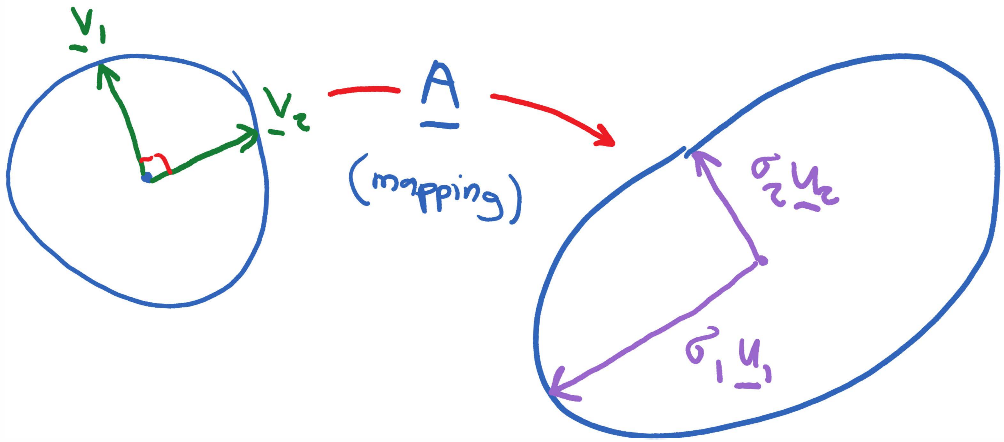
$\underline{v}$ is in the row space of $\underline{A}$

$\underline{u}$ is in the column space of $\underline{A}$

with $r = \text{rank}(\underline{A})$

What do σ, $\underline{u}$ + $\underline{v}$ represent?

Consider application of $\underline{A}$ to the unit circle



The singular decomposition gives the principal directions of the hyperellipses of $\underline{A}$ applied to the unit circle

Let $\underset{n\times r}{\hat{\underline{V}}} = [\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_r]$, $\underset{m\times r}{\hat{\underline{U}}} = [\underline{u}_1\ \underline{u}_2 \cdots \underline{u}_r]$

$$\underset{r\times r}{\hat{\underline{\Sigma}}} = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_r \end{bmatrix} \qquad r = \operatorname{rank}(\underline{A})$$

$$\underline{A}\,\hat{\underline{V}} = \hat{\underline{U}}\,\hat{\underline{\Sigma}}$$

where both $\hat{\underline{U}}$ and $\hat{\underline{V}}$ are unitary

$$\underset{r\times n}{\hat{\underline{V}}^T}\ \underset{n\times r}{\hat{\underline{V}}} = \underset{r\times r}{\underline{I}}, \quad \underset{r\times m}{\hat{\underline{U}}^T}\ \underset{m\times r}{\hat{\underline{U}}} = \underset{r\times r}{\underline{I}}$$

$\Rightarrow \underline{A} = \hat{\underline{U}}\,\hat{\underline{\Sigma}}\,\hat{\underline{V}}^T \leftarrow$ the reduced SVD

If $r < \min(m,n)$

$\Rightarrow$ Non-zero null space

$\Rightarrow$ There is a set of vectors that correspond to the singular values $\sigma = 0$

$$\underline{A}\,\underline{v} = \sigma\,\underline{u} = 0\,\underline{u} = \underline{0}$$

$\Rightarrow$ $\underline{v}$ is the null space of $\underline{A}$

The full SVD of $\underline{A}$ is then

$$\underset{m\times n}{\underline{A}} \Big[\overbrace{\underbrace{\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_r}_{\substack{r \text{ vectors in row space} \\ C(\underline{A}^T)}}}^{\hat{\underline{V}}}\ \underbrace{\underline{v}_{r+1} \cdots \underline{v}_n}_{\substack{n-r \text{ vectors in null space} \\ N(\underline{A})}} \Big] = \Big[\overbrace{\underbrace{\underline{u}_1 \cdots \underline{u}_r}_{\substack{r \text{ vectors in column space} \\ C(\underline{A})}}}^{\hat{\underline{U}}}\ \underbrace{\underline{u}_{r+1} \cdots \underline{u}_m}_{\substack{m-r \text{ vectors in left null space} \\ N(\underline{A}^T)}} \Big] \underset{m\times n}{\underline{\Sigma}}$$

$$\underbrace{\qquad\qquad\qquad}_{\underline{V}} \qquad \underbrace{\qquad\qquad\qquad}_{\underline{U}}$$

$\Rightarrow$ $\underset{m\times n}{\underline{A}} = \underset{m\times m}{\underline{U}}\ \underset{m\times n}{\underline{\Sigma}}\ \underset{n\times n}{\underline{V}^T}$ contains the orthonormal basis for all four matrix subspaces

Formal Definition

Let $\underline{A} \in \mathbb{R}^{m \times n}$ $\qquad m \geq n$ not required

also $\underline{A}$ might not be full rank

SVD of $\underline{A}$ is given by $\underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T$

$\underline{U} \in \mathbb{R}^{m \times m}$ is unitary

$\underline{V} \in \mathbb{R}^{n \times n}$ is unitary

$\underline{\Sigma} \in \mathbb{R}^{m \times n}$ is diagonal

$$\underline{U}^T \underline{A}\,\underline{V} = \underline{U}^T \underline{U}\,\underline{\Sigma}\,\underline{V}^T \underline{V}$$

$$\underline{U}^T \underline{A}\,\underline{V} = \underline{\Sigma}$$

diagonalization of $\underline{A}$

It is also assumed that all σ_j in $\underline{\Sigma}$ are real, non-negative and in non-increasing order

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0 \quad \text{for } p = \min(m,n)$$

To show real + non-negative consider $\underline{A}^T \underline{A}$

$$\underline{A}^T \underline{A} = (\underline{U}\,\underline{\Sigma}\,\underline{V}^T)^T (\underline{U}\,\underline{\Sigma}\,\underline{V}^T)$$

$$= \underline{V}\,\underline{\Sigma}^T \underline{U}^T \underline{U}\,\underline{\Sigma}\,\underline{V}^T$$

$= \underline{V}\,\underline{\Sigma}^2\,\underline{V}^T$

$\rightarrow$ Looks like an eigendecomposition of $\underline{A}^T\underline{A}$

Since $\underline{A}^T\underline{A}$ is normal $\Rightarrow$ $\underline{V}$ is unitary

Now consider $\underline{x}^T(\underline{A}^T\underline{A})\underline{x}$ for any $\underline{x}$

$$\underline{x}^T(\underline{A}^T\underline{A})\underline{x} = (\underline{A}\underline{x})^T(\underline{A}\underline{x}) = \underline{y}^T\underline{y} > 0$$

$\Rightarrow$ $\underline{A}^T\underline{A}$ is positive definite

$\Rightarrow$ $\underline{A}^T\underline{A}$ only has positive eigenvalues

Since $\underline{\Sigma}^2$ is the matrix of eigenvalues of $\underline{A}^T\underline{A}$

$\Rightarrow$ $\sigma_j = \sqrt{\lambda_j}$ $\Rightarrow$ will be positive + real

Theorem: Every matrix $\underline{A} \in \mathbb{R}^{m \times n}$ has an SVD and the singular values $\{\sigma_j\}$ are all uniquely determined

If $\underline{A}$ is square and all $\{\sigma_j\}$ are distinct, then $\{\underline{u}_j\}$ and $\{\underline{v}_j\}$ are uniquely determined up to a sign.

Properties:

Let $\underline{A} \in \mathbb{R}^{m \times n}$ with $p = \min(m,n)$

$r = \#$ of positive singular values

$r \leq p$

① $\text{rank}(\underline{A}) = r$

② $\text{range}(\underline{A}) = \text{span}(\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_r)$ (or column space)

$\text{null}(\underline{A}) = \text{span}(\underline{v}_{r+1}, \underline{v}_{r+2}, \ldots, \underline{v}_n)$ (or nullspace)

③ $\|\underline{A}\|_2 = \sigma_1$

$\|\underline{A}\|_F = (\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_r^2)^{1/2}$

Matrix norms based on SVD

④ Non-zero singular values of $\underline{A}$ are the square roots of the eigenvalues of

$\underline{A}^T \underline{A}$ or $\underline{A}\,\underline{A}^T$

⑤ If $\underline{A} = \underline{A}^T$, then σ_j is $|\lambda_j|$ of $\underline{A}$

⑥ If $\underline{A} \in \mathbb{R}^{m \times m}$, then $|\det(\underline{A})| = \prod_{i=1}^{m} \sigma_i$

$\Rightarrow$ If $\underline{A}$ is square, but one $\sigma_i = 0$,

then $\underline{A}^{-1}$ does not exist

Because of (2) above, the SVD says that any matrix can be made diagonal if one uses the proper row & column space basis

Consider $\underline{A}\,\underline{x} = \underline{b}$ $\qquad \underline{A} \in \mathbb{R}^{m \times n}$

$\underline{x} \in \mathbb{R}^n, \ \underline{b} \in \mathbb{R}^m$

$\underline{V}$ spans $\mathbb{R}^n$, while $\underline{U}$ spans $\mathbb{R}^m$

$\Rightarrow$ One can write $\underline{x}$ in terms of coordinates of $\underline{V}$

$$\underline{x}' = \underline{V}^T \underline{x} \quad \rightarrow \quad \underline{V}\,\underline{x}' = \underline{V}\,\underline{V}^T \underline{x}$$

$$\therefore \quad \underline{x} = \underline{V}\,\underline{x}'$$

Similarly

$$\underline{b}' = \underline{U}^T \underline{b}$$

$$\underline{A}\,\underline{x} = \underline{b} \Rightarrow \underline{U}^T \underline{A}\,\underline{x} = \underline{U}^T \underline{b}$$

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

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

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

$$\underline{\Sigma}\,\underline{x}' = \underline{b}'$$

Coordinates in $\underline{U}$

Diagonal matrix

Coordinates in $\underline{V}$

Uses of SVD:

1 Pseudo-Inverse

All matrices have $\underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T$

Define the pseudo-inverse as

$$\underline{A}^+ \underline{A} = \underline{I}, \quad \underline{A}\,\underline{A}^+ = \underline{I}$$

Note: $\underline{A}^{-1}$ might not exist

Let $\underline{A}^+ = \underline{V}\,\underline{\Sigma}^{-1}\underline{U}^T$ with

$$\underline{\Sigma}^{-1} = \begin{bmatrix} \sigma_1^{-1} & & 0 \\ & \sigma_2^{-1} & \\ 0 & & \ddots \end{bmatrix}$$

Recall for eigensystems:

$$\underline{A} = \underline{\Phi}\,\underline{\Lambda}\,\underline{\Phi}^T$$

$$\underline{A}^{-1} = \underline{\Phi}\,\underline{\Lambda}^{-1}\underline{\Phi}^T$$

$$A' = \underline{\Phi}\,\underline{\Lambda}\,\underline{\Phi}^T$$

Then

$$\underline{A}^+\underline{A} = (\underline{V}\,\underline{\Sigma}^{-1}\underline{U}^T)(\underline{U}\,\underline{\Sigma}\,\underline{V}^T)$$

$$= \underline{V}\,\underline{\Sigma}^{-1}\underline{\Sigma}\,\underline{V}^T = \underline{V}\,\underline{V}^T = \underline{I} \qquad (a)$$

$$\underline{A}\,\underline{A}^+ = (\underline{U}\,\underline{\Sigma}\,\underline{V}^T)(\underline{V}\,\underline{\Sigma}^{-1}\underline{U}^T)$$

$$= \underline{U}\,\underline{\Sigma}\,\underline{\Sigma}^{-1}\underline{U}^T = \underline{U}\,\underline{U}^T = \underline{I}$$

What happens if one or more $\sigma = 0$?

Then the corresponding diagonal elements in both $\underline{\Sigma}$ and $\underline{\Sigma}^{-1}$ are set to zero

Equation (a) no longer holds. Instead,

$$\underline{A}\,\underline{A}^+\underline{A} = \underline{A} \quad \& \quad \underline{A}^+\underline{A}\,\underline{A}^+ = \underline{A}^+$$

② Low Rank Approximations

Let

$$\underline{\Sigma}_j = \begin{bmatrix} 0 & & & & \\ & \ddots & & 0 & \\ & & 0 & & \\ & 0 & & \sigma_j & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}$$

Then

$$\underbrace{\underline{U}}_{m\times m}\ \underbrace{\underline{\Sigma}_j}_{m\times n}\ \underbrace{\underline{V}^T}_{n\times n} = \underbrace{\underbrace{\sigma_j}_{1\times 1}\ \underbrace{\underline{u}_j}_{m\times 1}\ \underbrace{\underline{v}_j^T}_{1\times n}}_{r=1}$$

with $\underline{u}_j$: jth column of $\underline{U}$

$\underline{v}_j$: jth column of $\underline{V}$

$$\Rightarrow \underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T = \sum_{j=1}^{r} \underline{U}\,\underline{\Sigma}_j\,\underline{V}^T$$

$$= \sum_{j=1}^{r} \sigma_j\, \underline{u}_j\, \underline{v}_j^T$$

$\Rightarrow$ Any matrix $\underline{A}$ can be written as the finite sum of rank 1 matrices

Theorem: Let $\underline{A}_\nu = \sum_{j=1}^{\nu} \sigma_j\, \underline{u}_j\, \underline{v}_j^T$ be a low rank approximation of $\underline{A}$, where $\nu \leq \text{rank}(\underline{A})$

Then, it can be shown that

$$\|\underline{A} - \underline{A}_\nu\|_2 = \inf_{\substack{\underline{B} \in \mathbb{R}^{m \times n} \\ \text{rank } \underline{B} \leq \nu}} \|\underline{A} - \underline{B}\|_2 = \sigma_{\nu+1}$$

Aside: inf (infimum)
greatest lower bound

where $\sigma_{\nu+1} = 0$, if $\nu = p = \min(m,n)$

$\Rightarrow$ $\underline{A}_\nu$ minimizes the error

One also can show that $\underline{A}_\nu$ minimizes the

$\|\underline{A} - \underline{A}_\nu\|_F$ error

To show this in an application, look at compression

with focus on gray-scale

An image is just a matrix with values

between 0 a d 255, with

0 = black, 255 = white

Let the image be represented by 256x512 pixels

Storing the full image takes

$256 \times 512 = 131072$ pixels (or data points)

Instead, store only the 5 largest singular values (i.e., $\nu = 5$)

$$\underset{\sim}{A} \rightarrow \text{Image} \approx \sigma_1 \underline{u}_1 \underline{v}_1^T + \sigma_2 \underline{u}_2 \underline{v}_2^T + \dots + \sigma_5 \underline{u}_5 \underline{v}_5^T$$

Size of compressed image

$$\underbrace{5}_{\nu \text{ values of } \sigma} + \underbrace{5}_{\nu}(\underset{\substack{u \\ 256\times 1}}{256} + \underset{\substack{v \\ 512\times 1}}{512}) = 3845 \text{ data points}$$

Compression ratio: $\dfrac{131072}{3845} \sim \dfrac{34}{1}$

Even more dramatic for larger images

Full image storage $\quad \mathcal{O}(mn)$

Compressed image storage $\quad \mathcal{O}(\nu(m+n))$