Review

Wednesday, March 22, 2017 9:05 AM

Observation $y = Ax + e$

unknown qty

error

observation matrix

assumed A is symmetric, positive-definite

Reconstruct / Estimate

$$\hat{x} = A^{-1} y = x + A^{-1} e$$

Best / Worst case reconstruction error

$$\frac{1}{\lambda_1^2} \|e\|_2^2 \le \|\hat{x} - x\|_2^2 \le \frac{1}{\lambda_N^2} \|e\|_2^2$$

equality is achieved for $\underline{e}$ pointing in the direction of $\underline{v}_N$

Average reconstruction error

assuming $e[n] \sim \text{Normal}(0, \sigma^2)$, $e[n]$ iid

$$E\left[\|\hat{x} - x\|_2^2\right] = \sigma^2 \sum_{n=1}^{N} \frac{1}{\lambda_n^2}$$

Singular Value Decomposition

Wednesday, March 22, 2017 9:15 AM

We generally are interested in non-sym+def matrices.

$$y = Ax, \quad y \in \mathbb{R}^M, \quad A \text{ is } M \times N, \quad x \in \mathbb{R}^N$$

The singular value decomposition (SVD) of A is

$$A = U \Sigma V^T$$

$U = [u_1 | u_2 | \cdots | u_R]$ is an $M \times R$ matrix

$u_m \in \mathbb{R}^M$ are orthogonal and normalized so that

$U^T U = I$. Note $UU^T \neq I$ when $R < M$ in general.

$\{u_m\}$ form an orthobasis for the range space of A

$V = [v_1 | v_2 | \cdots | v_R]$ is an $N \times R$ matrix

$v_n \in \mathbb{R}^N$ and are orthonormal

$V^T V = I$ but, in general, $VV^T \neq I$ when $R < N$

$\{v_n\}$ form an orthobasis for the range space of A^T

$\rightarrow$ Range(A^T) consists of everything which is orthogonal to the null space of A.

Σ is an $R \times R$ diagonal matrix with positive entries

$$\Sigma = \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_R \end{bmatrix}$$

σ_r are the singular values of A

$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_R$

SVD properties

Wednesday, March 22, 2017 9:29 AM

$$A = U \Sigma V^T$$

$$\underbrace{A^T A}_{} = V \Sigma \underbrace{U^T U}_{I} \Sigma V^T = V \Sigma^2 V^T$$

symmetric, positive semi-definite (some eigen values could be zero)

σ_r's are the square roots of the eigenvalues of A^TA (actually of the non-zero eigenvalues of A^TA)

$$A A^T = U \Sigma^2 U^T$$

The non-zero eigenvalues of A^TA and AA^T are the same so the σ_r are also the square roots of the eigenvalues of AA^T

$A \in \mathbb{R}^{M \times N}$

if $M \geq N$, then R is the number of linearly independent columns

if $M \leq N$, then R is the number of linearly independent rows

A is full rank if $R = \min(M, N)$

$$\big[A\big] = \underbrace{\big[U\big]}_{6\times 3} \underbrace{\big[\Sigma\big]}_{3\times 3} \underbrace{\big[V^T\big]}_{4\times 3}$$

$M > N$

$y = Ax$ is overdetermined

x is 4×1

y is 6×1

$$\big[A\big] = \big[U\big]\big[\Sigma\big]\big[V^T\big]$$

$M < N$, $y = Ax$ is under-determined

$$\big[A\big] = \big[U\big]\big[\Sigma\big]\big[V^T\big]$$

$M = N = R$ full-rank

A is square

SVD and Least Squares

Wednesday, March 22, 2017 9:44 AM

Since $A = U \Sigma V^T = \left[u_1 \,|\, u_2 \,|\, \cdots \,|\, u_R \right] \begin{bmatrix} \sigma_1 & & 0 \\ & \sigma_2 & \\ 0 & & \ddots \, \sigma_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_R^T \end{bmatrix}$

$$A = \sum_{r=1}^{R} \sigma_r u_r v_r^T$$

"Solving" $y = Ax$ using SVD

$y \in \mathbb{R}^M$ $\quad A \in \mathbb{R}^{M \times N}$

$x \in \mathbb{R}^N$

Given y, we want to find x in such a way that

1. when the unique solution exists, use that
2. when no solution exists, return something reasonable
3. when there are an infinite number of solutions, choose the "best" one.

Define a residual $\quad$ Least-Squares framework

$r = y - Ax \quad \Rightarrow$ optimize $\min_{x \in \mathbb{R}^N} \|y - Ax\|_2^2$

Solution

let $x = V\underline{\alpha} + V_0 \underline{\alpha}_0 = \left[V \,|\, V_0 \right] \begin{bmatrix} \underline{\alpha} \\ \underline{\alpha}_0 \end{bmatrix}$

V is $N \times R$ and from the SVD $(A = U \Sigma V^T)$

V_0 is $N \times (N-R)$ and the colums of V_0 are orthonormal and orthogonal to the columns of V

$V^T V = I, \quad V_0^T V_0 = I \quad V^T V_0 = 0$

$$\left[V \,|\, V_0 \right]^T \left[V \,|\, V_0 \right] = I_{N \times N}$$

$\alpha = V^T x \quad , \quad \alpha_0 = V_0^T x$

SVD and Least Squares cont.

Wednesday, March 22, 2017 9:58 AM

$$\alpha = V^T x \quad , \quad \alpha_0 = V_0^T x$$

Note, $\quad x = V\alpha + V_0\alpha_0 = VV^T x + V_0 V_0^T x$

$$= \left[VV^T + V_0 V_0^T \right] x = x$$

$$y = U\beta + U_0 \beta_0$$

$$U^T U = I \quad , \quad U_0^T U_0 = I \quad , \quad U^T U_0 = 0$$

$$[U \mid U_0]^T [U \mid U_0] = I_{m \times m}$$

$$\beta = U^T y \quad , \quad \beta_0 = U_0^T y$$

$$r = y - Ax$$

$$r = U\beta + U_0\beta_0 - U\Sigma V^T (V\alpha + V_0\alpha_0)$$

$$= U\beta + U_0\beta_0 - U\Sigma\alpha$$

$$= \underbrace{U_0\beta_0 + U(\beta - \Sigma\underline{\alpha})}$$

We want to choose $\underline{\alpha}$ that minimizes $\|r\|_2^2$

$$\|r\|_2^2 = \langle \cdot \, , \cdot \rangle = \langle U_0\beta_0 , U_0\beta_0 \rangle - 2\cancel{\langle U_0\beta_0 , U(\beta - \Sigma\alpha) \rangle} \quad 0$$

$$+ \langle U(\beta - \Sigma\alpha), U(\beta - \Sigma\alpha) \rangle$$

has a $U_0^T U$ factor which is 0

$$= \|U_0\beta_0\|_2^2 + \underbrace{\|U(\beta - \Sigma\underline{\alpha})\|_2^2}_{\text{minimized when } \hat{\alpha} = \Sigma^{-1}\beta}$$

The x which minimizes $\|r\|_2^2$ is

$$\hat{x} = V\hat{\alpha} = V\Sigma^{-1}\beta = \underbrace{V\Sigma^{-1}U^T}\, y$$

$$A = U\Sigma V^T \qquad A \cdot V\Sigma^{-1}U^T = U\Sigma \overbrace{\cancel{V^T V}}^{I_{R\times R}} \underbrace{\cancel{\Sigma\Sigma^{-1}}}_{I_{R\times R}} U^T$$

$$\hat{x} = UU^T y$$

Psuedo Inverse

Wednesday, March 22, 2017 10:12 AM

If $A = \underbrace{U \Sigma V^T}_{\text{SVD of } A}$ Then the psuedo-inverse of A

$$A^{\dagger} = V \Sigma^{-1} U^T$$

example from when you were a child

$$y = Ax$$

$$A^{\dagger} = (A^T A)^{-1} A^T$$ ← only applies to full-rank A, and A has $M > N$

and another form ...

$$A^T (A A^T)^{-1}$$ ← (I think)

1. why $y \in \operatorname{span}(\{u_1, \dots, u_m\})$ then $\beta_0 = U_0^T y = 0$ and then $r = 0$.

2. If $R < N$ the solution is not unique. In this case V_0 has at least one column, and any part of a vector x in the range of V_0 is "not seen by A" since

$$A V_0 \alpha_0 = U \Sigma \underbrace{V^T V_0}_{=0} \alpha_0 = 0$$

ex: $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ 0 & & & 0 \end{bmatrix}$

it won't matter

$x = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$

what d is, A cannot see it.

define $x' = \hat{x} + V_0 \alpha_0$

$\left(\text{ex: } V_0 = [0\ 0\ 0\ 1]^T \right)$

for any $\alpha_0 \in \mathbb{R}^{N-R}$ we have the same r since $A x' = A \hat{x}$

represents all possible solutions by choice α_0

$$\|x'\|_2^2 = \langle \hat{x} + V_0 \alpha_0, \hat{x} + V_0 \alpha_0 \rangle$$

$$= \langle \hat{x}, \hat{x} \rangle + 2 \langle \hat{x}, V_0 \alpha_0 \rangle + \langle V_0 \alpha_0, V_0 \alpha_0 \rangle$$

$$= \|\hat{x}\|_2^2 + 2 \langle V \Sigma^{-1} U^T y, V_0 \alpha_0 \rangle + \|\alpha_0\|_2^2$$

cause this to go to 0

$$= \|\hat{x}\|_2^2 + \|\alpha_0\|_2^2$$

SVD Least-Squares Solution

Wednesday, March 22, 2017 10:27 AM

Using the SVD to find the solution as described on the previous pages minimizes both the residual **and** the norm of the solution!

in Matlab

```
x = A\y;
```