Overview

SVD

Solve Square Linear Systems

Can the SVD $A = U\Sigma V^T$ be used to solve *square* linear systems?
At what cost (once the SVD is known)?

$$A\vec{x} = \vec{b}$$

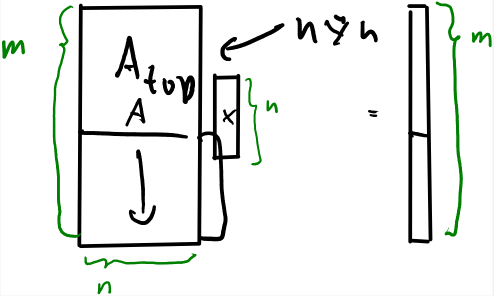
$$U\Sigma V^T \vec{x} = \vec{b}$$

$$\vec{x} = \underbrace{V \underbrace{\Sigma^{-1} \underbrace{U^T \vec{b}}_{n^2}}_{n}}_{n^2}$$

$$\begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sigma_1 x_1 \\ \vdots \\ \sigma_n x_n \end{pmatrix}$$

Tall and Skinny Systems

Consider a ‘tall and skinny’ linear system, i.e. one that has more equations than unknowns:



In the figure: $m > n$. How could we solve that?

"overdetermined"

$A_{top}x = b_{top}$

Instead of $Ax=b$, ask for

"Which $\vec{x}$ makes $\|Ax-b\|_2^2$ as small as possible?"

$\vec{r} = A\vec{x}-\vec{b}$ "residual"

$\|Ax-b\|_2^2 = \|r\|_2^2 = \underleftarrow{r_1^2 + \cdots + r_m^2}$

↳ "least squares"

Solving Least-Squares

How can I actually *solve* a least-squares problem $Ax \cong \boldsymbol{b}$?

Suppose Q is orthogonal: $\|Q\vec{y}\|_2 \overset{\otimes}{=} \|y\|$

$$\min_x \|Ax-b\|_2^2 = \|U\Sigma V^T\vec{x}' - \vec{b}\|_2^2$$

$$\overset{\otimes}{=} \|\underbrace{U^T(U}_{I}\Sigma V^T\vec{x} - \vec{b})\|_2^2 = \|\Sigma \underbrace{V^T\vec{x}}_{y' = V^T\vec{x}} - U^T b\|_2^2$$

$$= \min_y \|\Sigma\vec{y} - U^T\vec{b}\|_2^2$$

$$= \left\| \begin{pmatrix} \overset{\checkmark}{\sigma_1} & & \\ & \ddots & \\ & & \overset{\checkmark}{\sigma_n} \end{pmatrix} \begin{matrix} \overset{\times}{y_1} \\ \vdots \\ y_n \end{matrix} - \overbrace{\overset{\checkmark}{U^T} \overset{\checkmark}{\vec{b}}}^{\vec{z}} \right\|$$

$$= (\sigma_1 \overset{\downarrow}{y_1} - z_1)^2 + \cdots (\sigma_n \overset{\downarrow}{y_n} - z_n)^2$$

$$|\sigma_1| \geq |\sigma_2| \cdots \geq |\sigma_n| \geq 0$$

$$\sigma_1 \cdots \sigma_k > 0 \qquad \sigma_{k+1} \cdots \sigma_n = 0$$

$$= (\sigma_1 y_1 - z_1)^2 + \cdots + (\sigma_k y_k - z_k)^2$$

$$\left.\begin{aligned} &+ \left(\overset{0}{\cancel{\sigma_{k+1}}} y_{k+1} - z_{k+1}\right)^2 \\ &\cdots \\ &+ \left(\overset{0}{\cancel{\sigma_n}} y_n - z_n\right)^2 \end{aligned}\right] \begin{aligned} &z_{k+1}^2 \\ &+ \cdots \\ &+ z_n^2 \end{aligned}$$

$$y_i = \frac{z_i}{\sigma_i} \Rightarrow (\sigma_i y_i - z_i)^2$$

$$= \left(\cancel{\sigma_i} \frac{z_i}{\cancel{\sigma_i}} - x_i\right)^2 = 0$$

$$\min \|\Sigma \vec{y} - U^T b\|$$

$$y_i = \begin{cases} (U^T b)_i / \sigma_i & \sigma_i \neq 0 \\ 0 & \sigma_i = 0 \end{cases}$$

doesn't matter, pick

$$\Sigma^{-1} = \begin{pmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & \boxed{1/\sigma_n} \end{pmatrix}$$

not allowe

$$\Sigma^{+} = \begin{pmatrix} 1/\sigma_1 & & & & & \\ & \ddots & & & & \\ & & 1/\sigma_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

„pseudo inverse of a diagonal“

with that:

$$\vec{y} = \Sigma^{+} U^{T} \vec{b}$$

Still wanted: $\|Ax - b\|_2^2$

$$\vec{y} = V^{T} \vec{x} \iff V\vec{y} = \vec{x}$$

full & skinny $\longrightarrow$ $\vec{x} = V \Sigma^{+} U^{T} \vec{b}$

square & inv. $\longrightarrow$ $\boxed{\vec{x} = V \Sigma^{-1} U^{T} b}$

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

↑

pseudoinv. of A

$(x = A^{-1}b)$

$$\vec{x} = A^{+}\underline{b}$$

↑

solves $\min \|Ax-b\|_2^2$

$$A^T A x = A^T b$$

↑ normal equations

$$\mathrm{cond}(A^T A) \approx \mathrm{cond}(A^T)\,\mathrm{cond}(A)$$

$$\approx \mathrm{cond}(A)^2$$

$$(10^6)^2$$

In-class activity: SVD and Least Squares

The Pseudoinverse: A Shortcut for Least Squares

How could the solution process for $A\boldsymbol{x} \cong \boldsymbol{b}$ be with an SVD$A = U\Sigma V^T$ be ‘packaged up’?

The Normal Equations

You may have learned the 'normal equations' $A^T A\boldsymbol{x} = A^T\boldsymbol{b}$ to solve $A\boldsymbol{x} \cong \boldsymbol{b}$.
Why not use those?

$n+1$

$$p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n$$

$$V \begin{pmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix}$$

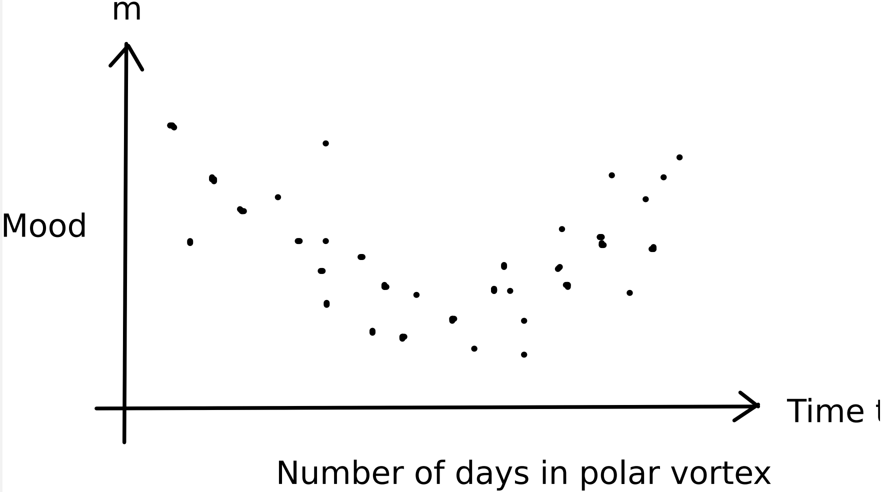
$$p(x_0) = y_0$$

$$p(x_n) = y_n$$

Outline

Fitting a Model to Data

How can I fit a model to measurements? E.g.:



Demo: Data Fitting using Least Squares

Outline