LECTURE 10 HOUSEHOLDER REFLECTORS QR

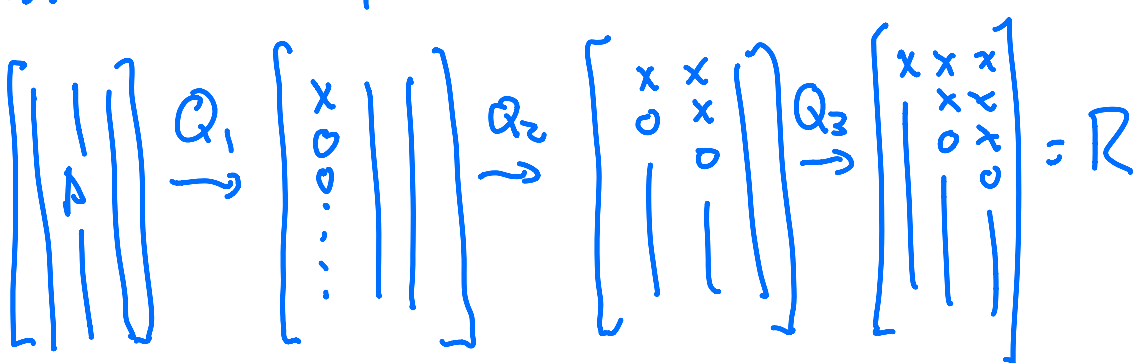
Previously: GS amounts to the successive application of upper triangular matrices R_i:

$$A \underbrace{R_1 R_2 R_3 \cdots R_n}_{R^{-1}\ \text{(upper triangular)}} = Q$$

An alternative QR: Householder. We create Q_i's so that

$$\underbrace{Q_n Q_{n-1} \cdots Q_2 Q_1}_{Q^*} A = R \quad \text{, since}$$

Q_i's are unitary $\Rightarrow Q$ is unitary.



$Q_1 A$ $\quad$ $Q_2 Q_1 A$ $\quad$ $Q_3 Q_2 Q_1 A$

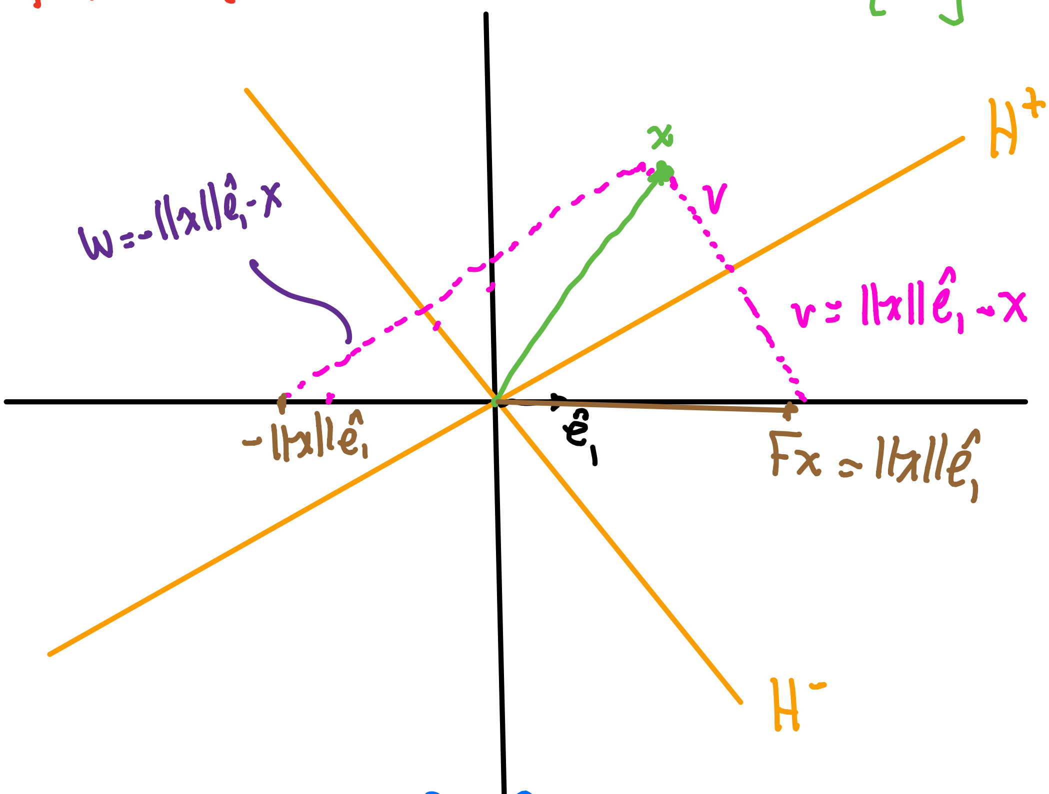
Q_k introduces zeros below the k^{th} row in the k^{th} column w/o destroying the zeros the zeros from before

Each

$$Q_k = \begin{array}{c c} \overset{k-1}{\longleftrightarrow} & \\ \begin{bmatrix} I \updownarrow_{k-1} & 0 \\ 0 & F \end{bmatrix} & \updownarrow m-k+1 \\ & \underset{m-k+1}{\longleftrightarrow} \end{array}$$

F is the Householder Reflector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} \quad (e_1)$$



$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} \xrightarrow{F} \begin{bmatrix} \|x\| \\ 0 \\ 0 \\ \vdots \end{bmatrix} = \|x\|\,\hat{e}_1$$

To find the reflection about H^+:

$$Py = y - \frac{vv^*}{v^*v}y = y - \left(\frac{v^*y}{vv^*}\right)v$$

for some $y \in \mathbb{C}^m$

is the projection of some vector y onto H^+.

To get to the $\hat{e}_1$ axis:

$$Fy = \left(I - \frac{2vv^*}{v^*v}\right)y = y - 2v\left(\frac{v^*y}{v^*v}\right)$$

$\therefore$ Matrix $F = I - \dfrac{2vv^*}{v^*v}$

Rmk: the projector P (rank $m-1$) and the reflector F (full rank & unitary) only differ by the presence of the 2!

The Better of 2 Reflectors:

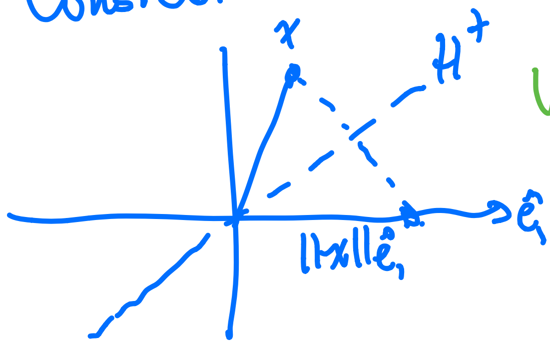
The vector $x \xrightarrow{F} z\|x\|\hat{e}_1 \quad z \in \mathbb{C}$ w/ $|z| = 1$

$\therefore$ a circle of possible reflections

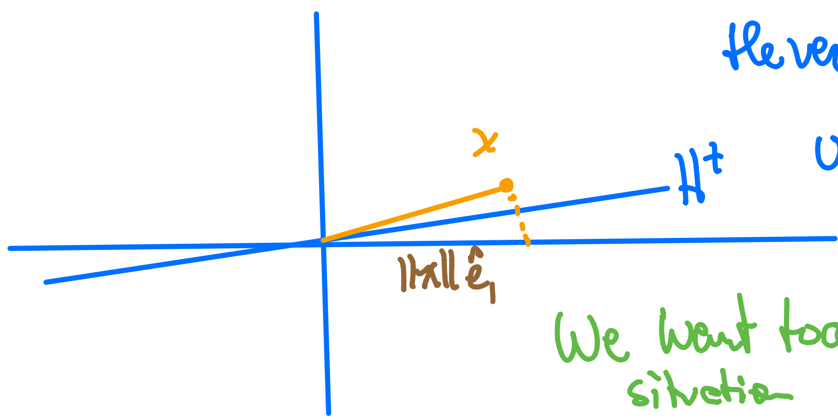
If $z \in \mathbb{R}$ then $z = \pm 1$ lead to H^+ & H^- cases, either is OK, but choice is numerically driven.

We want to avoid subtracting nearby quantities.

Consider



We like this situation



the vector $x \sim \|x\|\hat{e}_1$

$v = \|x\|\hat{e}_1 - x \sim 0$

We want to avoid this situation

$\therefore$ For this second case why not reflect about H^-? (see figure above)

Choose $v = -\text{sgn}(x_1)\|x\|\hat{e}_1 - x$

(if $x_1 = 0 \Rightarrow \text{sgn} = 0$)

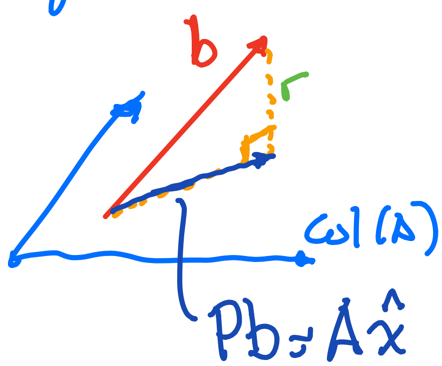
then $\|v\|$ is never smaller than $\|x\|$ //

See Algorithm 10.1
Householder QR
Cost $\sim O(mn^2)$ $\sim O(n^3)$ if square.
but more stable numerically.

Lecture 11 LSQ (least squares)

Solving $Ax=b$

$r = Ax - b$



if $r=0 \Rightarrow Ax=b$

$A\hat{x} = Pb$, $b = Pb + r$

Choose r to be "shortest"

$\|r\|^2 = \|b - Pb\|^2 = \|(I-P)b\|^2$

the minimizer will be $\hat{x}$. $r \perp \mathrm{col}(A)$

The orthogonality condition

$$A^* r = 0$$

or $$A^*(b - A\hat{x}) = 0$$

or $$A^*A\hat{x} = A^*b$$ The Normal Equations

solving for $\hat{x} = (A^*A)^{-1}A^*b$

which satisfies $A\hat{x} = Pb = \underbrace{A(A^*A)^{-1}A^*}_{P}b$ //

ex) $Ax = b$ $\quad A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix}$ $\quad b = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$

$\quad\quad\quad\quad\quad\quad\quad\quad$ $\shortparallel \;\; \shortparallel$

$\quad\quad\quad\quad\quad\quad\quad\quad$ $a_1 \;\; a_2$

so clear that $b \neq x_1 a_1 + x_2 a_2$

but $Pb = \hat{x}_1 a_1 + \hat{x}_2 a_2$ can be found.

so let $r = b - Pb = (I - P)b$

Find $\hat{x}$ s.t. $A^T r = 0$

$A^T A \hat{x} = A^T b$ the Normal Eqs

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 5 \\ 5 & 13 \end{bmatrix}$$

$$\hat{x} = (A^T A)^{-1} A^T b = \begin{bmatrix} 13 & -5 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 3 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$Pb = A\hat{x} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix}$$

$$b - Pb = r = \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix} = (I-P)b \quad \therefore \|r\|_2 = 6 \quad //$$

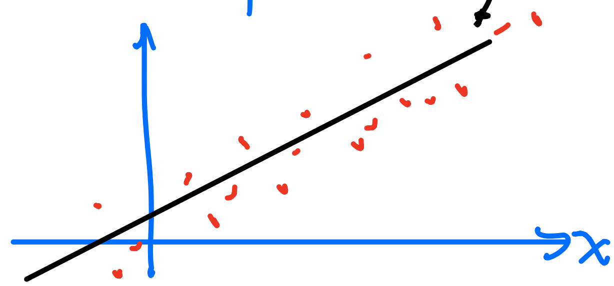
FITTING DATA TO MODELS

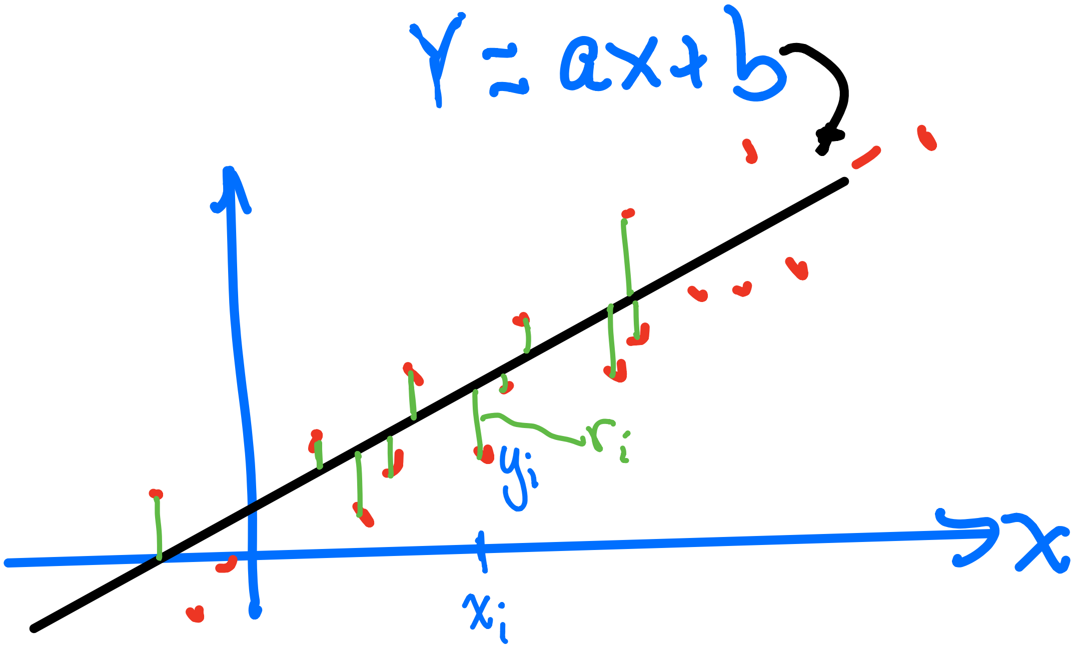
Data:

x_1	y_1
x_2	y_2
$\vdots$	$\vdots$
x_m	y_m

"Model" look like a straight line

$Y = ax + b$





$r_i = Y(x_i) - y_i$

Variance
energy $= \sum |r_i|^2$

Find a, b s.t. energy is minimized.

A least squares problem:

let $A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}$ $\quad z = \begin{bmatrix} a \\ b \end{bmatrix}$ $\quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$

$r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}$

$Az = y + r$

Solve for $\hat{z}$ s.t. $\|r\|_2^2$ is smallest. //

POLYNOMIAL DATA FITTING

$\{x_i\}_{i=1}^m \quad \{y_i\}_{i=1}^m \qquad x \in \mathbb{C}^m$

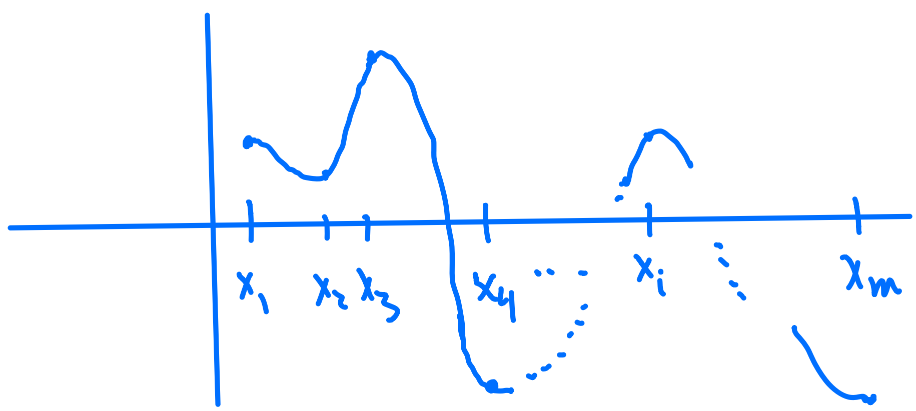
$y \in \mathbb{C}^m$

$\exists\, !$ polynomial interpolant

$p(x) = c_0 + c_1 x + c_2 x^2 + \dots\, c_{m-1} x^{m-1}$

s.t. at x_i $\;p(x_i) = y_i$

m eqs & m unknowns, the powers of x are linearly independent.



the condition $p(x_i) = y_i \quad i = 1, 2, \dots, m$

generates

Rmk: in interpolation we insist that $p(x_i) = y_i,\ i = 1 \dots m$

$\therefore$ we need m equations for m unknowns.

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & & & & \\ 1 & x_m & x_m^2 & \cdots & x_m^{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

Vandermonde Matrix

So imagine fitting 10^6 points... using a polynomial of deg $10^6 - 1$??

... Highly unstable process on a finite precision machine. //

Rmk: in approximation we don't insist that $p(x_i) = y_i$ for $i = 1, \ldots m$. Instead we propose an approximating polynomial that has some optimizing characteristic. For example, we can insist that we find an approximating polynomial that minimizing a residual $p(x_i) - y_i$. //