Honors Numerical Analysis

Lecture 11

Recall from last time that computing $A = QR$ using Householder reflections constructs a sequence of $\hat{Q}_1^*, \ldots, \hat{Q}_n^*$ such that $\hat{Q}_n^* \cdots \hat{Q}_1^* A = R$.

① $\hat{Q}_1^*$ was determined so that $\hat{Q}_1^* \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} = \begin{pmatrix} \|\vec{a}_1\| \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

② Find $\hat{Q}_2^*$ so that $\hat{Q}_2^* \hat{Q}_1^* \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ \vdots & \vdots \\ a_{m1} & a_{m2} \end{pmatrix} = \begin{pmatrix} \|\vec{a}_1\| & \tilde{a}_{12} \\ 0 & \|\tilde{a}_2\| \\ \vdots & 0 \\ 0 & \vdots \\ & 0 \end{pmatrix}$

where $\|\tilde{a}_2\| = \sqrt{\sum_{i=2}^{m} |\tilde{a}_{i2}|^2}$. This can easily be achieved by combining a Householder reflection with an Identity matrix:

$$\hat{Q}_2^* = \left(\begin{array}{c|ccc} I_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & R_{m-1} & \\ 0 & & & \end{array}\right)$$

where R_{m-1} is $(m-1) \times (m-1)$

and $R_{m-1} \begin{pmatrix} \tilde{a}_{22} \\ \tilde{a}_{32} \\ \vdots \\ \tilde{a}_{m2} \end{pmatrix} = \begin{pmatrix} \|\tilde{a}_2\| \\ 0 \\ \vdots \\ 0 \end{pmatrix}$

③ Each $\hat{Q}_j^*$ can similarly be constructed as

$$\hat{Q}_j = \left(\begin{array}{c|c} I_j & 0 \\ \hline 0 & R_{m-j} \end{array}\right).$$

Truncated QR

Suppose that $A = QR$ is truncated so that

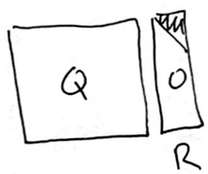
$\tilde{Q} \in \mathbb{C}^{m \times k}$, $\tilde{R} \in \mathbb{C}^{k \times n}$ ~~and~~ .

$\hookrightarrow = Q(:, 1:k)$ $\searrow = R(1:k, :)$

Hopefully $A \approx \tilde{Q}\tilde{R}$, depending on k.

What is $\|A - \tilde{Q}\tilde{R}\|$?

Consider a full QR of A: $A =$



with $Q(:, 1:k) = \tilde{Q}$. Then

$$\|A - \tilde{Q}\hat{R}\| = \|QR - \hat{Q}\tilde{R}\| = \|R - Q^* \cancel{R}\hat{Q}\hat{R}\|$$

$$= \left\|R - \begin{pmatrix} I_k \\ 0 \end{pmatrix}\tilde{R}\right\|$$

$$= \left\|R - \begin{pmatrix} \tilde{R} \\ 0 \end{pmatrix}\right\|$$

$$= \left\| \begin{array}{cl} \begin{pmatrix} 0 \\ \hline \blacktriangledown \\ \hline 0 \end{pmatrix} & \begin{array}{l} k \\ k+1 \\ n \\ \downarrow m \end{array} \\ n & \end{array} \right\|$$

Applications: <u>Linear Regression</u>

Given data: (x_i^1, x_i^2, y_i) for $i = 1 \dots n$

Generative model: $y = a + bx_1 + cx_2 + \epsilon$

$\epsilon \sim N(0, \sigma^2)$ error or measurement noise.

You observe $y_i = a + bx_i^1 + cx_i^2 + \epsilon_i$

y_i: dependent variable

x_i^1, x_i^2: independent variables

Assumption is that noise only appears in y_i.

Given an estimate of a, b, c: $\hat{a}, \hat{b}, \hat{c}$, then the residuals are given as

$$r_i = \underset{\text{observed value}}{y_i} - \underbrace{(\hat{a} + \hat{b}x_i^1 + \hat{c}x_i^2)}_{\text{predicted value, given estimates } \hat{a}, \hat{b}, \hat{c}.}$$

Question: How do we estimate a, b, c?

One option: minimize the squared residuals:

$$\min_{a,b,c} \|\vec{r}\|_2^2 = \min_{a,b,c} \sum_{i=1}^n (y_i - a - bx_i^1 - cx_i^2)^2$$

This is a least squares problem.

Form the least square system:

$$\min_{a,b,c} \left\| \underbrace{\begin{pmatrix} 1 & x_1^1 & x_1^2 \\ 1 & x_2^1 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n^1 & x_n^2 \end{pmatrix}}_{X} \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{\vec{a}} - \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}}_{\vec{y}} \right\| = \min_{\vec{a}} \| \underset{\text{design matrix}}{X} \vec{a} - \vec{y} \|$$

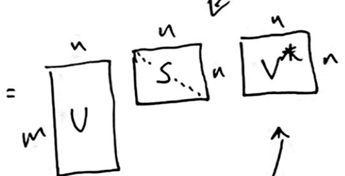
The Singular Value Decomposition

Let $A \in \mathbb{C}^{m \times n}$, $m > n$, and rank $A = n$.

Then A can be factored as:

positive scaling

$A = U S V^*$



rotation, reflection

with U, V orthogonal / unitary

S diagonal with entries

$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$

Mathematically, the SVD can be computed by observing that A^*A is positive semi-definite and therefore can be written as: $A^*A = V S^2 V^*$ ← an eigen decomposition.

Then, set $U = A V S^{-1}$ $\Rightarrow$ $A = U S V^*$.

This is numerically more unstable than dealing with A directly since $\kappa(A^*A) = \underline{\kappa(A)^2}$.

Instead, observe that the matrix (assume $m=n$)

$$H = \begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix} \in \mathbb{C}^{2m \times 2m}$$ is Hermitian,

and that if $A = USV^*$ then

$$\underbrace{H \begin{pmatrix} V & V \\ U & -U \end{pmatrix} = \begin{pmatrix} V & V \\ U & -U \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & -S \end{pmatrix}}_{\text{eigendecomposition of } H.}$$

$\Rightarrow$ No matrix squaring, but deals with $2m \times 2m$ matrix instead of an $m \times n$ one...

More on how to efficiently construct an SVD later on...

Pseudo-inverse If $m > n$, the A is not invertible, but:

Define the pseudo-inverse of A to be

$$A^+ = \underset{m \times n}{V}\, \underset{n \times n}{S^{-1}}\, \underset{n \times n}{U^*}$$

Even though A is not square,

$$A^+ A = (VS^{-1}U^*)(USV^*)$$

$$= \underbrace{V \underbrace{S^{-1} \underbrace{U^* U}_{I} S}_{I} V^*}_{I}$$

$= I.$

Applications to least squares problems.