One more SVD-related issue.

Solving least-squares problems:

compute solution x to $Ax=b$

for an overdetermined ($A \in \mathbb{C}^{m\times n}$, $m>n$)

Can't always make equality true, so instead:

compute x s.t.

$$\|Ax-b\|_2 = \min_{y\in\mathbb{C}^n} \|Ay-b\|_2$$

Here: we'll use SVD to compute a solution

We'll assume: $m \geq n$, and $\operatorname{rank}(A)=n$.

$A = U\Sigma V^*$, U, V unitary

$= \tilde{U}\tilde{\Sigma}\tilde{V}^*$ (reduced SVD)

$V=\tilde{V}$, $\Sigma = \begin{pmatrix} \tilde{\Sigma} \\ 0 \end{pmatrix}$ an $m\times n$ matrix

$\tilde{\Sigma}$: $n\times n$ and invertible.

$U = (\tilde{U} \,\vdots\, \hat{U})$, $\tilde{U}$: $m\times n$

$\hat{U}$: $m\times(m-n)$

Solve: $\min_y \|Ay-b\|_2 = \min_y \|Ay-b\|_2^2$

$$= \min_y \|U^*(Ay-b)\|_2^2$$

$$= \min_y \|\Sigma V^* y - U^* b\|_2^2$$

$$= \min_y \left\| \begin{pmatrix} \tilde{\Sigma} V^* y - \tilde{U}^* b \\ 0 - \hat{U}^* b \end{pmatrix} \right\|_2^2$$

Pythagoras $\rightarrow$ $= \min_y \left(\|\underbrace{\tilde{\Sigma}}_{\text{invertible}} V^* y - \tilde{U}^* b\|_2^2 + \underbrace{\|\hat{U}^* b\|_2^2}_{\text{independent of } y!} \right)$

choose $y = V\tilde{\Sigma}^{-1} U^* b$

$\longrightarrow$ $\min_y \|Ay-b\|_2^2 = \min_y \|\hat{U}^* b\|_2^2 = \boxed{\|\hat{U}^* b\|_2^2}$

$x = y = V\tilde{\Sigma}^{-1} U^* b$ solves least-squares problem (uniquely!)

$\boxed{(A^*A)x = A^*b}$ (normal eqns).

The QR decomposition

MATH 6610 Lecture 11

September 28, 2020

Trefethen & Bau: Lecture 7

Orthogonalization

The main goal of orthogonalization:

Given $\{a_j\}_{j=1}^n \subset \mathbb{C}^m$, compute $\{q_j\}_{j=1}^n$ such that:

$$\langle q_j, q_k \rangle = \delta_{j,k}, \qquad \text{span}\{a_1, \ldots, a_n\} = \text{span}\{q_1, \ldots, q_n\}$$

Actually want

$\text{span}\{a_1 \ldots a_j\} = \text{span}\{q_1 \ldots q_j\}$

$\forall j = 1, 2 \ldots n.$

Orthogonalization

The main goal of orthogonalization:

Given $\{a_j\}_{j=1}^n \subset \mathbb{C}^m$, compute $\{q_j\}_{j=1}^n$ such that:

$$\langle q_j, q_k \rangle = \delta_{j,k}, \qquad \mathrm{span}\{a_1, \ldots, a_n\} = \mathrm{span}\{q_1, \ldots, q_n\}$$

Why?

One reason is that the $\mathbb{C}^m$-orthogonal projector onto $\mathrm{span}\{a_1, \ldots, a_n\}$ is given by,

$$P = QQ^*, \qquad Q = \begin{pmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{pmatrix}$$

Second reason: this is actually how eigenvalues are computed (!!).

Gram-Schmidt orthogonalization

One essentially explicit algorithm to orthogonalize is Gram-Schmidt.

Input: n vectors $\{a_j\}_{j=1}^n$. (Assume they're linearly independent for now.)
Output: n vectors $\{q_j\}_{j=1}^n$.

Define: $P_k := \sum_{j=1}^{k} q_j q_j^*$ (this is basically QQ^*)

orthogonal projector onto span$\{q_1 \ldots q_k\}$.

Define $P_0 := 0$.

For $j = 1 \ldots n$

$u_j = a_j - P_{j-1} a_j$ ($u_j \neq 0$ since a_j are LI).

$q_j = u_j / \|u_j\|$.

end

Gram-Schmidt to QR

input G-S output

$$\{a_1, \ldots, a_n\} \longrightarrow \{q_1, \ldots, q_n\},$$

$$\operatorname{span}\{a_1, \ldots, a_j\} = \operatorname{span}\{q_1, \ldots, q_j\} \ (1 \leqslant j \leqslant n)$$

$$\langle q_i, q_k \rangle = \delta_{i,k}$$

Gram-Schmidt to QR

$$\{a_1, \ldots, a_n\} \quad \longrightarrow \quad \{q_1, \ldots, q_n\},$$
$$\operatorname{span}\{a_1, \ldots, a_j\} = \operatorname{span}\{q_1, \ldots, q_j\} \;\; (1 \leqslant j \leqslant n)$$

We can rewrite this to explicitly express the original vectors a_j in terms of the orthogonalized vectors q_j.

$$\begin{cases} u_j = a_j - P_{j-1} a_j = a_j - \sum_{k=1}^{j-1} q_k (q_k^* a_j) \\ u_j / \|u_j\| = q_j \end{cases}$$

$$a_j = u_j + \sum_{k=1}^{j-1} q_k (q_k^* a_j) = \|u_j\| q_j + \sum_{k=1}^{j-1} q_k (q_k^* a_j)$$

define $q_k^* a_j = r_{\substack{\cancel{jk} \\ kj}}$ (scalar)

$\|u_j\| = r_{jj}$

$$a_j = \sum_{k=1}^{j} q_k \, r_{\substack{\cancel{jk} \\ kj}} \qquad (\star)$$

$$\begin{pmatrix} | \\ a_j \\ | \end{pmatrix} = \begin{pmatrix} | & & | \\ q_1 & \cdots & q_j \\ | & & | \end{pmatrix} \begin{pmatrix} r_{1j} \\ \vdots \\ r_{jj} \end{pmatrix} \qquad (\forall j = 1 \ldots n)$$

$$\underset{\text{"}A\text{"}}{\begin{pmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{pmatrix}} = \underset{\text{"}Q\text{"}}{\begin{pmatrix} | & & | \\ q_1 & \cdots & q_n \\ | & & | \end{pmatrix}} \underset{\text{upper-triangular} \quad \text{"}R\text{"}}{\begin{pmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & & r_{2n} \\ 0 & 0 & r_{33} & & \vdots \\ \vdots & \vdots & 0 & \ddots & \vdots \\ & & \vdots & & r_{nn} \end{pmatrix}}$$

can make Q unitary:

$$A = \begin{pmatrix} | & & | & | & & | \\ q_1 & \cdots & q_n & q_{n+1} & \cdots & q_m \\ | & & | & | & & | \end{pmatrix} \begin{pmatrix} \overset{n}{\begin{matrix} \diagdown & \\ 0 & \end{matrix}} \; n \\ \hdashline 0 \end{pmatrix}$$

any
$\mathbb{C}^m$ completion
(orthonormal)

is unitary!

is upper-triangular.

The QR decomposition

In fact, these computations implies the following result:

Theorem

Let $A \in \mathbb{C}^{m\times n}$ be any matrix. Then there exists a unitary matrix $Q \in \mathbb{C}^{m\times m}$, and an upper-triangular matrix $R \in \mathbb{C}^{m\times n}$ such that

$$A = QR.$$

If A has full rank, then the diagonal entries of R can be chosen to be positive.

think about Gram-Schmidt

$(r_{jj} = \|u_j\|)$

If A is rank-deficient, $m \geq n$:

$$A = QR$$

$$\begin{pmatrix} | & | & | & \cdots & | \end{pmatrix} = \begin{pmatrix} | & | & | & | & \cdots & | & | \end{pmatrix}\begin{pmatrix} x & x & x & x & \\ 0 & x & x & x & \\ \vdots & 0 & 0 & x & \ddots \\ \vdots & \vdots & \vdots & 0 & \\ 0 & 0 & 0 & 0 & \end{pmatrix}$$

still is upper-triangular, with zero elements in some diagonal entries.

$n < m$: