Householder Reflectors

~~Modified Gram-Schmidt~~

MATH 6610 Lecture 13

October 2, 2020

Trefethen & Bau: Lecture 10 ✓

Orthogonalization

The main goal of orthogonalization:

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

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

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

$1 \le j \le n$

Orthogonalization

The main goal of orthogonalization:

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

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

Any algorithm to accomplish this (e.g., Gram-Schmidt) implies:

$$A = QR, \quad A = \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{pmatrix}, \quad Q = \begin{pmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{pmatrix},$$

with R upper triangular.

Orthogonalization

The main goal of orthogonalization:

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

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

Any algorithm to accomplish this (e.g., Gram-Schmidt) implies:

$$A = QR, \quad A = \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{pmatrix}, \quad Q = \begin{pmatrix} | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \end{pmatrix},$$

with R upper triangular.

We've seen "classical" (unstable) Gram-Schmidt and "modified" Gram-Schmidt.

((cf. demos)

Householder reflectors

Let P be an orthogonal projection matrix. Then $I - 2P$ is Hermitian, unitary, and involutory.

$P \in \mathbb{C}^{n \times n}$

$P^2 = P$, $P^* = P$

Def'n: A matrix A is involutory (i.e. is an involution) if

$A^{-1} = A$.

$(AA = I)$

$I - 2P$ is unitary + Hermitian: HW

$I - 2P$ is involutory: direct computation

Def'n: A Householder reflector is any matrix $I - 2P$, where P is an orthogonal projector.

Householder reflectors

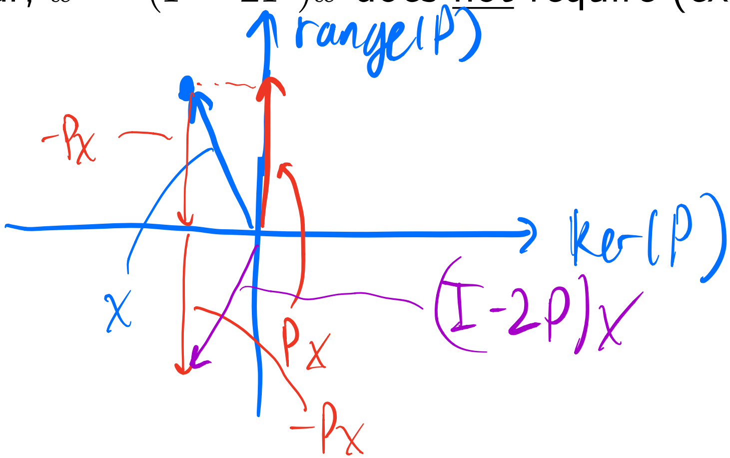
Let P be an orthogonal projection matrix. Then $I - 2P$ is Hermitian, unitary, and involutory.

Thus, application of this matrix, $x \mapsto (I - 2P)x$, is well-conditioned.

In particular, if P is a rank-1 projector, then there is a unit vector v such that

$$P = vv^*. \quad (\|v\|_2 = 1)$$

(And in particular, $x \mapsto (I - 2P)x$ does <u>not</u> require (expensive) matrix-vector multiplications.)



↓r $(I-2P)x = (I-2vv^*)x = x - 2v(\underbrace{v^*x}_{1\ \text{IP}})$

vector subtraction.

Use of Householder reflectors

Our main use of these reflectors is the following:

Given $x \in \mathbb{C}^m$, we want to achieve:

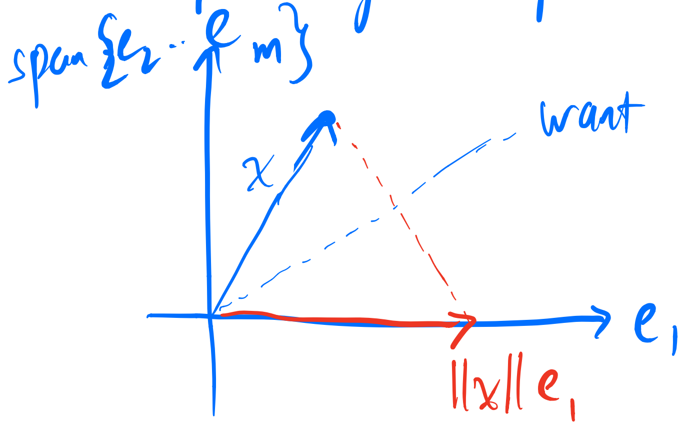
$$x \xrightarrow{\text{Householder reflector}} \overbrace{\|x\| e^{i\theta}}^{\text{scalar}} e_1,$$

for some $\theta \in [0, 2\cancel{,\pi})$.

$2\pi)$

$e_1 = (1, 0, 0, \ldots 0)^* \in \mathbb{C}^m$

(For now: keep things simple and assume $\theta = 0$)



want HR that reflects around this axis.

We want to do this w/ a rank-1 projector $P = vv^*$.

Want: $(I - 2P)x = \|x\| e_1$ (assuming $\theta = 0$)

$(I - 2vv^*)x = \|x\| e_1$

$\frac{x - \|x\| e_1}{2v^* x} = v \longrightarrow$ this is a vector that points in the direction $x - \|x\| e_1$.

Use of Householder reflectors

Our main use of these reflectors is the following:

Given $x \in \mathbb{C}^m$, we want to achieve:

$$x \quad \xrightarrow{\text{Householder reflector}} \quad \|x\| e^{i\theta} e_1,$$

for some $\theta \in [0, 2, \pi)$.

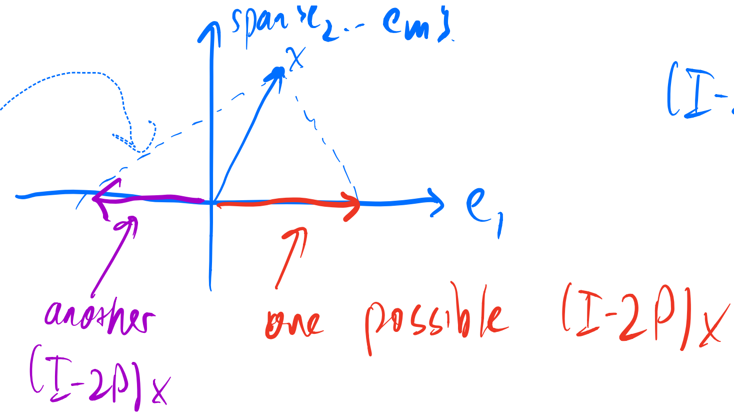
This is achieved by the reflector $I - 2vv^*$, with v given by

$$v = \frac{x - \|x\| e^{i\theta} e_1}{\|x - \|x\| e^{i\theta} e_1\|},$$

for arbitrary θ.

Also: want this to be numerically stable.

Empirically: want this reflector to change x by as large an amount as possible.



$$(I-2P)x = e_1 \|x\| e^{i\theta}$$

$(\theta = 0)$

$(\theta = \pi)$

choice: dictated by choice of θ;

If the above 2 choices were the only options for θ: choose second one

Use of Householder reflectors

Our main use of these reflectors is the following:

Given $x \in \mathbb{C}^m$, we want to achieve:

$$x \quad \xrightarrow{\text{Householder reflector}} \quad \|x\|e^{i\theta}e_1,$$

for some $\theta \in [0, 2, \pi)$.

This is achieved by the reflector $I - 2vv^*$, with v given by

$$v = \frac{x - \|x\|e^{i\theta}e_1}{\|x - \|x\|e^{i\theta}e_1\|}, \qquad (\star, \star)$$

for arbitrary θ.

$HR = I - 2vv^*$

For numerical stability, this reflector should make large changes to x, rather than small changes.

$x = (x_1, x_2, \dots x_m)^T$

Largest change achieved by selecting

$$e^{i\theta} = -\frac{x_1}{|x_1|}. \qquad (\star)$$

Use of Householder reflectors, II

We now have the following procedure:
Given $x \in \mathbb{C}^m$, we compute $v \in \mathbb{C}^m$ such that

input

$$(I - 2P)x = (I - vv^*)x = c\, e_1,$$

2

for some scalar $c \in C$.

Use of Householder reflectors, II

We now have the following procedure:
Given $x \in \mathbb{C}^m$, we compute $v \in \mathbb{C}^m$ such that

$$(I - 2P)x = (I - vv^*)x = c\, e_1,$$

for some scalar $c \in C$.

Put another way: we can, via an efficiently-applicable unitary transform, map x to e_1.

the main realization

A new viewpoint on QR

Given a matrix A, all versions of Gram-Schmidt perform operations associated with a triangular matrix R^{-1},

$$A \longrightarrow AR^{-1} = Q$$

Thus, this is "triangular orthogonalization".

orthogonalization (of A) by means of triangular (linear-algebraic) operations.

A new viewpoint on QR

Given a matrix A, all versions of Gram-Schmidt perform operations associated with a triangular matrix R^{-1},

$$A \longrightarrow AR^{-1} = Q$$

Thus, this is "triangular orthogonalization".

We can use Householder reflectors to instead perform:

$$A \longrightarrow Q^*A = R,$$

which is an "orthogonal triangularization".

How?

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

1.) Generate a HR that transforms first column of A to (a multiple of) e_1.

$Q_1 = (I - 2v_1v_1^*)$ (HR)

$$Q_1 A = \begin{pmatrix} x & x & x & \cdots & x \\ 0 & x & x & \cdots & x \\ 0 & x & x & \ddots & x \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x & x & \cdots & x \end{pmatrix} \quad (\star)$$

2.) Generate a HR in $\mathbb{C}^{m-1}$ that reflects the $(m-1)$-vector in column 2 corresponding to truncating first element. $(\star)$

$Q_2 = (I - 2v_2v_2^*)$ (HR) $Q_2 \in \mathbb{C}^{(m-1)\times(m-1)}$

$\tilde{Q}_2 = \begin{pmatrix} 1 & 0 \\ 0 & Q_2 \end{pmatrix} \in \mathbb{C}^{m\times m}$ and is unitary.

$$\tilde{Q}_2 Q_1 A = \begin{pmatrix} x & x & x & x & \cdots & x \\ 0 & x & x & x & \cdots & \vdots \\ \vdots & 0 & x & \vdots & \ddots & x \\ 0 & 0 & x & \cdots & \cdots & x \end{pmatrix}$$

Continue...

$$\underbrace{\tilde{Q}_{m-1}\tilde{Q}_{m-2}\cdots\tilde{Q}_2 Q_1}_{Q^*} A = \overbrace{\begin{pmatrix} x & x & x & x & \cdots & x \\ 0 & x & x & & & \vdots \\ \vdots & 0 & x & & \ddots & \vdots \\ \vdots & \vdots & 0 & & & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & x \end{pmatrix}}^{R}$$

A new viewpoint on QR

Given a matrix A, all versions of Gram-Schmidt perform operations associated with a triangular matrix R^{-1},

$$A \longrightarrow AR^{-1} = Q$$

Thus, this is "triangular orthogonalization".

We can use Householder reflectors to instead perform:

$$A \longrightarrow Q^* A = R,$$

which is an "orthogonal triangularization".

We expect the latter to be more stable since we are simply applying unitary (well-conditioned) matrices to A.

(This is in fact what many implementations of QR decompositions use.)

There is one more orthogonalization strategy: Givens' Rotations