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:

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

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.

Householder reflectors

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

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

$$P = vv^*.$$

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

Use of Householder reflectors

Our main use of these reflectors is the following:

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

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

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

Use of Householder reflectors

Our main use of these reflectors is the following:

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

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

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.

Use of Householder reflectors

Our main use of these reflectors is the following:

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

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

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.

For numerical stability, this reflector should make large changes to x, rather than small changes.
Largest change achieved by selecting

$$e^{i\theta} = -\frac{x_1}{|x_1|}.$$

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 $c \in C$.

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.

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".

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".

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.)