Lecture 5

One Dimensional Projection Methods

L. Olson

Department of Computer Science
University of Illinois at Urbana-Champaign

January 25, 2007

Projection Template Algorithm 5.1

while $\|r\| > tol$
    Generate $V = [v_1 \ldots v_m]$ of $\mathcal{K}$
    Generate $W = [v_1 \ldots v_m]$ of $\mathcal{L}$
    $r = b - Ax$
    $y = x + V(W^T AV)^{-1} W^T r$
end

One dimensional Projection Template Algorithm 5.1.5

Let $\mathcal{K} = \text{span}\{v\}$ and $\mathcal{K} = \text{span}\{w\}$

```
Given x, A, b
r = b − Ax
p = Av
while ||r|| > tol
    p = Av
    α = (r, w)/(p, w)
    y = x + αv
    r = r − αp
end
```

- one matvec per iteration
- now pick v and w

SD

Let $v = r$ and $w = r$

```
Given x, A s.p.d., b
r = b − Ax
p = Ar
while ||r|| > tol
    p = Ar
    α = (r, r)/(p, r)
    y = x + αr
    r = r − αp
end
```

- after each step $r_{k+1} \perp w = r_k$

MR

Let $v = r$ and $w = Ar$

```
Given x, A, b
r = b − Ax
p = Ar
while ‖r‖ > tol
    p = Ar
    α = (r,p)/(p,p)
    y = x + αr
    r = r − αp
end
```

- after each step $r_{k+1} \perp w = Ar_k$
- A indefinite is valid, but $p.d.$ needed for convergence

RNSD

Let $v = A^T r$ and $w = Av$

```
Given x, A, b
r = b - Ax
while ||r|| > tol
    v = A^T r
    w = A * v
    α = (v,v)/(w,w)
    y = x + αv
    r = r - αAv
end
```

- after each step $r_{k+1} \perp w = AA^T r_k$
- SD on $A^T A x = A^T b$
- convergence even if A is indefinite (and nonsingular)

Improvement?

Consider updates

$$x_{k+1} = x_k + \alpha_k r_k$$

The residual update is

$$\begin{aligned} r_{k+1} &= r_k - \alpha_k A r_k \\ &= (I - \alpha_k A)^k r_0 \\ &= P_k(A) r_0 \end{aligned}$$

So we're reducing the residual with a polynomial in A. Presumably

$$x = A^{-1}b \leftarrow x_0 + P_k(A) r_0 \text{as } k \rightarrow \infty$$

New iteration

So we want to construct x_m from

$$A^{-1}b \approx x_m = x_0 + P_{m-1}(A)r_0$$

Letting $r_0 = \sum_{j=1}^{n} \gamma_j v_j$, with (λ_i, v_i) the eigenpairs of A we have:

$$r_i = P_i(A)r_0 = \sum_{j=1}^{n} \gamma_j P_j(\lambda_j) v_j$$

Error reduction depends on P's ability to reduce initial error components

How to pick P_i?

Our construction so far shows P_i of the form

$$P_i(A) = \prod_{j=1}^{i} (I - \alpha_j A)$$

- one approach: let α_i be the Chebyshev zeros. This is *Chebyshev* iteration.
- problem: P_i completely changes form as i increases.

Notice:

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k r_k \\ &= x_{k-1} + \alpha_{k-1} r_{k-1} + \alpha_k r_k \\ &= x_0 + \alpha_0 r_0 + \alpha_1 r_1 \ldots \alpha_k r_k \\ &\in \text{span}\{r_0, Ar_0, A^2 r_0, \ldots, A^k r_0\} \end{aligned}$$

since $r_{k+1} = (I - \alpha_k A)^k r_0$.

Krylov

Krylov Space

An m-dimensional subspace of the form

$$\mathcal{K}_m(A; v) \equiv= \{v, Av, A^2v, \dots, A^{m-1}v\}$$

is called a *Krylov subspace*

The goal then is to find the “best” x_m from the Krylov subspace generated by A and r_0: $\mathcal{K}_m(A; r_0)$.

i.e. Construct better polynomials from $\mathcal{K}_m$

A basis for $\mathcal{K}_m$

If we're going to use $\mathcal{K}_m$, we will need a suitable basis.

- the basis $r_0, Ar_0, A^2r_0, \ldots A^{m-1}r_0$ is not good
- $A^k r_0$ approach the dominant eigenvector as k increases
- the basis vectors become nearly linearly dependent (think monomials)
- really want to look for an orthogonal basis...

An orthogonal basis

Many ways to generate an orthogonal basis:

1. Gram-Schmidt (GS): simple
2. modified Gram-Schmidt (MGS): numerically better than GS
3. modified Gram-Schmidt with reorthogonalization (MGSR): maintains
4. Householder: more stable than MGS, but more expensive. Orthogonal vectors not available until the end (fine for some methods).
5. Givens: parallelizability

recall classical Gram-Schmidt

```
Given {x_1, dots, x_r} linearly independent.
r_11 = ||x_1||_2
if r_11 = 0, return
q_1 = x_1/r_11
for j = 2, dots, r
    r_ij = (x_j, q_i) for i = 1, 2, ..., j-1
    q_j = x_j - \sum_{i=1}^{i=j-1} r_ij q_i
    r_jj = ||q_j||_2
    if rjj = 0, return
    q_j = q_j/r_jj
end
```

Now, $x_j = \sum_{i=1}^{j} r_{ij} q_i$, and we get $X = QR$.

recall modified Gram-Schmidt

```
Given {x_1, dots, x_r} linearly independent.
r_{11} = ||x_1||_2
if r_{11} = 0, return
q_1 = x_1 / r_{11}
for j = 2, dots, r
    q_j = x_j
    for i = 1, ..., j-1
        r_{ij} = (q_j, q_i)
        q_j = q_j - r_{ij} q_i
    end
    r_{jj} = ||q_j||_2
    if r_j j = 0, return
    q_j = q_j / r_{jj}
end
```

With reorthogonalization, simply reapply MGS to the q set.

Householder

Gram-Schmidt:

$$A \underbrace{R_1 R_2 \dots R_n}_{R^{-1}} = Q$$

R_j are upper triangular (so is R^{-1}. Then $A = QR$

Householder:

$$\underbrace{Q_n \dots Q_2 Q_1}_{Q=\text{unitary}}$$

So $A = QR$.

Difference: applying sequence of unitary matrices (based on reflections) instead of a sequence of triangular matrices.

Householder

Pick Q_k so that it zeros the kth subcolumn, leaving all previous columns the same:

$$A = \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{Q_1} Q_1 A = \begin{bmatrix} \times' & \times' & \times' \\ & \times' & \times' \\ & \times' & \times' \\ & \times' & \times' \\ & \times' & \times' \end{bmatrix}$$

$$\xrightarrow{Q_2} Q_2 Q_1 A = \begin{bmatrix} \times' & \times' & \times' \\ & \times'' & \times'' \\ & & \times'' \\ & & \times'' \\ & & \times'' \end{bmatrix}$$

$$\xrightarrow{Q_3} Q_3 Q_2 Q_1 A = \begin{bmatrix} \times' & \times' & \times' \\ & \times'' & \times'' \\ & & \times''' \\ & & \\ & & \end{bmatrix}$$

Projecting

Let P_u pjoject orthogonally onto u:

$$P_u = \frac{uu^T}{u^T u}$$

and

$$P_u x = u \frac{(u, x)}{||u||^2}$$

since x project orthogonally onto u, $((I - P_u)x, x) = 0$.

- $(I - P_u)x$ is the difference between x and $P_u x$:

$$(I - P_u)x = x - P_u x$$

- the projection onto u minus the difference, gives a reflection

$$y = P_u x - (I - P_u)x = (2P_u - I)x$$

Householder

Choose Q_k of the form

$$Q_k \begin{bmatrix} I & 0 \\ 0 & F \end{bmatrix}$$

I is $(k-1) \times (k-1)$ and F is $(m-k+1) \times (m-k+1)$. F is a Householder reflector. F should satisfy

$$\begin{bmatrix} \times \\ \times \\ \times \\ \times \\ \times \end{bmatrix} \xrightarrow{F} \begin{bmatrix} ||x|| \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$