

CS475 / CM375

Lecture 9: Oct 11, 2011

Conjugate Gradient
Readings: [TB] Chapt 38, p. 293-301
[Saad] Sect 5.3 and 6.7

CS475/CM375 (c) 2011 P. Poupart & J. Wan

1

Method of Conjugate Directions

- Each new direction is “orthogonal” in some sense to previous search directions
- It turns out the best strategy is to require A -orthogonality (or conjugacy)
 - i.e., p, q are A -orthogonal if $p^T A q = 0$
- Definition: the A -inner product is defined as:
$$(p, q)_A \equiv p^T A q$$
- The A -norm is defined as $\|p\|_A = \sqrt{(p, p)_A}$

CS475/CM375 (c) 2011 P. Poupart & J. Wan

2

Gram-Schmidt Process

- Construct a set of orthogonal vectors
- Suppose the previous search directions $p^0, p^1, \dots, p^{k-1}$ are A -orthogonal. Given the current r^k , construct p^k
- Let $p^k = r^k + \sum_{i=0}^{k-1} \beta_i p^i$ then

$$(p^k, p^j)_A = 0 \Rightarrow (r^k, p^j)_A + (\sum_{i=0}^{k-1} \beta_i p^i, p^j)_A = 0$$

$$(r^k, p^j)_A + \beta_j (p^j, p^j)_A = 0$$

$$\beta_j = -\frac{(r^k, p^j)_A}{(p^j, p^j)_A}$$

CS475/CM375 (c) 2011 P. Poupart & J. Wan

3

Conjugate Gradient Picture

CS475/CM375 (c) 2011 P. Poupart & J. Wan

4

Conjugate Gradient Method

- Construct a set of A -orthogonal search vectors $\{p^k\}$ by the residual vectors $\{r^k\}$

$$\text{-- i.e., } p^k = r^k - \sum_{i=0}^{k-1} \frac{(r^k, p^i)_A}{(p^i, p^i)_A} p^i$$

CS475/CM375 (c) 2011 P. Poupart & J. Wan

5

Conjugate Gradient (CG) I

x^0 = initial guess; $r^0 = b - Ax^0$

for $k = 0, 1, \dots, n - 1$

compute $p^k = r^k - \sum_{i=0}^{k-1} \frac{(r^k, p^i)_A}{(p^i, p^i)_A} p^i$

$$x^{k+1} = x^k + \alpha_k p^k$$

$$r^{k+1} = r^k - \alpha_k A p^k$$

end

Notes:

- $\alpha_k = (r^k, p^k) / (p^k, p^k)_A$
- $r^{k+1} = b - Ax^{k+1}$

CS475/CM375 (c) 2011 P. Poupart & J. Wan

6

Useful Facts

- Krylov subspace

$$\begin{aligned}
 \text{span}\{p^0, \dots, p^{k-1}\} &= \text{span}\{r^0, \dots, r^{k-1}\} \\
 &= \text{span}\{r^0, Ar^0, \dots, A^{k-1}r^0\} \\
 &\equiv K_k(A, r^0) \\
 &= k\text{-dim Krylov subspace}
 \end{aligned}$$
- Since $r^k \perp \text{span}\{r^0, r^1, \dots, r^{k-1}\}$
 - i.e. $(r^k, r^j) = 0 \quad j = 0, 1, \dots, k-1$
 Then $r^k \perp \text{span}\{p^0, \dots, p^{k-1}\}$

CS475/CM375 (c) 2011 P. Poupart & J. Wan

7

Useful Facts

- $(r^k, p^k) = (r^k, r^k)$
 - Proof:

CS475/CM375 (c) 2011 P. Poupart & J. Wan

8

Useful Facts

- $(r^k, p^i)_A = 0 \quad i = 0, 1, \dots, k-2$
 – Proof:

CS475/CM375 (c) 2011 P. Poupart & J. Wan

9

Speed up

- By using the previous facts, we can speed up the computation of p^k by simplifying the summation
- $$p^k = r^k - \frac{(r^k, p^{k-1})_A}{(p^{k-1}, p^{k-1})_A} p^{k-1}$$

 – Proof:

CS475/CM375 (c) 2011 P. Poupart & J. Wan

10

Speed up

- Speed up the computation of $(r^k, p^{k-1})_A$ by avoiding any matrix multiplication
- $(r^k, p^{k-1})_A = -\frac{1}{\alpha_{k-1}}(r^k, r^k)$
 - Proof:

CS475/CM375 (c) 2011 P. Poupart & J. Wan

11

Speed up

- Similarly, speed up the computation of $(p^{k-1}, p^{k-1})_A$ by avoiding any matrix multiplication
- $(p^{k-1}, p^{k-1})_A = \frac{1}{\alpha_{k-1}}(r^{k-1}, r^{k-1})$
 - Proof:

CS475/CM375 (c) 2011 P. Poupart & J. Wan

12

Efficient Conjugate Gradient (CG)

```

 $x^0$  = initial guess;  $r^0 = b - Ax^0$ 
for  $k = 0, 1, \dots, n-1$ 
     $\beta_{k-1} = (r^k, r^k) / (r^{k-1}, r^{k-1})$  ( $\beta_{-1} = 0$ )
     $p^k = r^k + \beta_{k-1} p^{k-1}$ 
     $\alpha_k = (r^k, r^k) / (p^k, Ap^k)$ 
     $x^{k+1} = x^k + \alpha_k p^k$ 
     $r^{k+1} = r^k - \alpha_k Ap^k$ 
end

```

Notes

1. Only 1 matrix-vector multiplication, 2 inner products
2. At most n A -orth vectors in $\mathfrak{R}^n$. Terminate in at most n steps $\rightarrow$ exact solution