Lecture-11

Rate of Convergence of CG

Algorithm 5.2

Given x_0;

set $r_0 \leftarrow Ax_0 - b,\ \ p_0 \leftarrow -r_0, k \leftarrow 0$

While $r_k \neq 0$

$$a_k \leftarrow -\frac{r_k^T r_k}{p_k^T A p_k};$$

$$x_{k+1} \leftarrow x_k + a_k p_k;$$

$$r_{k+1} \leftarrow r_k + a_k A p_k;$$

$$b_{k+1} \leftarrow \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k};$$

$$p_{k+1} \leftarrow -r_{k+1} + b_{k+1} p_k;$$

$$k \leftarrow k+1;$$

end (*while*)

We only need to know values of x, p and r only for 2 iterations.

Major computations: matrix-vector product, two inner products, and three vector sums.

Key points

- According to theorem 5.3 Algorithm 5.2 should converge at most n steps.
- Convergence less than n iterations, depending on the eigenvalues of matrix A.
- If A dose not have favorable eigenvalues, then precondition A to get faster convergence.

Theorem 5.4

If A has only r distinct eigenvalues, then the CG iteration will terminate at the solution in at most r iterations.

Main points

Want to show:

$$\| x_{k+1} - x^* \|_A^2 \le \min_{P_k} \max_{1 \le i \le n} \left[1 + l_i P_k(l_i)\right]^2 \| x_0 - x^* \|_A^2$$

Use this: (Theorem 5.3)

Define polynomial $P_k^*(A) = g_0 I + g_1 A + \ldots + g_k A^k$

Use orthogonal eigenvectors n_i of A.

Show n_i are also eigenvectors of $P_k(A)$

Rate of Convergence

$$x_{k+1} = x_0 + a_0 p_0 + \ldots + a_k p_k$$

By construction

$$x_{k+1} = x_0 + g_0 r_0 + g_1 A r_0 + \ldots + g_k A^k r_0$$

(Theorem 5.3)

Define polynomial:

$$P_k^*(A) = g_0 I + g_1 A + \ldots + g_k A^k$$

Therefore

$$x_{k+1} = x_0 + P_k^*(A) r_0 \qquad \text{(D)}$$

Now

$$\frac{1}{2} \| x - x^* \|_A^2 = f(x) - f(x^*)$$

$$\| z \|_A^2 = z^T A z$$

$$f(x) = \frac{1}{2} x^T A x - b^T x$$

Rate of Convergence

$$\frac{1}{2}\| x-x^* \|_A^2 = \frac{1}{2}(x-x^*)^T A(x-x^*)$$
$$= \frac{1}{2}(x^T - x^{*T})(Ax - Ax^*)$$
$$= \frac{1}{2}x^T Ax - \frac{1}{2}x^{*T}Ax - \frac{1}{2}x^T Ax^* + \frac{1}{2}x^{*T}Ax^*$$
$$= \frac{1}{2}x^T Ax - \frac{1}{2}b^T x - \frac{1}{2}x^T b + \frac{1}{2}x^{*T}Ax^*$$
$$= \frac{1}{2}x^T Ax - \frac{1}{2}b^T x - \frac{1}{2}x^T b + x^{*T}Ax^* - \frac{1}{2}x^{*T}Ax^*$$
$$= \frac{1}{2}x^T Ax - \frac{1}{2}b^T x - \frac{1}{2}x^T b + x^{*T}b - \frac{1}{2}x^{*T}Ax^*$$
$$= \frac{1}{2}x^T Ax - b^T x - (\frac{1}{2}x^{*T}Ax^* - b^T x^*)$$
$$= f(x) - f(x^*)$$

$$\| z \|_A^2 = z^T Az$$

$$f(x) = \frac{1}{2}x^T Ax - b^T x$$

$$\frac{1}{2}\| x-x^* \|_A^2 = \frac{1}{2}(x-x^*)^T A(x-x^*) = f(x) - f(x^*)$$

$$x_{k+1} = x_0 + a_0 p_0 + \ldots + a_k p_k$$

By construction

$$f(x) = \frac{1}{2}x^T Ax - b^T x$$

According to Theorem 5.2 x_{k+1} minimizes f, hence $\| x-x^* \|_A^2$

Or

$\| x_0 + P_k^*(A)r_0 - x^* \|_A^2$ $\quad x_{k+1} = x_0 + P_k^*(A)r_0$ From (D)

Therefore, P_k^* solves the following problem:

$$\min_{P_k} \| x_0 + P_k(A)r_o - x^* \|_A$$

We know $r_0 = Ax_0 - b = Ax_0 - Ax^* = A(x_0 - x^*)$

$$x_{k+1} = x_0 + P_k^*(A) r_0$$

From (D)

$$\begin{aligned} x_{k+1} - x^* &= x_0 + P_k^*(A) r_0 - x^* \\ &= (x_0 - x^*) + P_k^*(A) r_0 \\ &= (x_0 - x^*) + P_k^*(A) A (x_0 - x^*) \\ &= [I + P_k^*(A) A](x_0 - x^*) \end{aligned} \qquad \text{(A)}$$

Assume $n_i,\ l_i$ are eigenvectors & eigenvalues of A

$$x_0 - x^* = \sum_{i=l}^{n} x_i v_i$$

Show are also eigenvectors of

$$P_k^*(A) = g_0 I + g_1 A + \ldots + g_k A^k$$

$$P_k(A) n_i = g_0 I n_i + g_1 A n_i + g_2 A^2 n + \ldots + g_k A^k n_i$$

$$P_k(A) n_i = g_0 n_i + g_1 l_i n_i + g_2 l_i A n_i \ldots + g_k A^{k-1} l_i n_i$$

$$P_k(A) n_i = g_0 n_i + g_1 l_i n_i + g_2 l_i^2 n_i \ldots + g_k A^{k-2} l_i^2 n_i$$

$$P_k(A) n_i = g_0 n_i + g_1 l_i n_i + g_2 l_i^2 n_i + \ldots + g_k l_i^k n_i$$

$$P_k(A) n_i = (g_0 + g_1 l_i + g_2 l_i^2 \ldots + g_k l_i^k) n_i$$

$$P_k(A) n_i = P(l_i) n_i$$

Therefore

We know $x_0 - x^* = \sum_{i=l}^{n} x_i v_i$

$$x_{k+1} - x^* = [I + P_k^*(A)A](x_0 - x^*) \qquad \text{From (A)}$$

$$x_{k+1} - x^* = \sum_{i=l}^{n} [I + P_k^*(A)A] x_i v_i$$

$$x_{k+1} - x^* = \sum_{i=l}^{n} [x_i v_i + P_k^*(A) A x_i v_i]$$

$$x_{k+1} - x^* = \sum_{i=l}^{n} [x_i v_i + P_k^*(A)\, l_i x_i v_i]$$

$$x_{k+1} - x^* = \sum_{i=l}^{n} [x_i v_i + l_i P_k^*(l_i) x_i v_i]$$

$$x_{k+1} - x^* = \sum_{i=l}^{n} [1 + l_i P_k^*(l_i)] x_i v_i$$

$$x_{k+1} - x^* = \sum_{i=l}^{n} [1 + l_i P_k^*(l_i)] x_i v_i$$

$$\| x_{k+1} - x^* \|_A^2 = \left(\sum_{i=l}^{n} [1 + l_i P_k^*(l_i)] x_i n_i^T \right) A \left(\sum_{i=l}^{n} [1 + l_i P_k^*(l_i)] x_i n_i \right)$$

$$\| x_{k+1} - x^* \|_A^2 = \left(\sum_{i=l}^{n} [1 + l_i P_k^*(l_i)] x_i n_i^T \right) \left(\sum_{i=l}^{n} {}_i[1 + l_i P_k^*(l_i)] x_i A n_i \right)$$

$$\| x_{k+1} - x^* \|_A^2 = \left(\sum_{i=l}^{n} [1 + l_i P_k^*(l_i)] x_i n_i^T \right) \left(\sum_{i=l}^{n} [1 + l_i P_k^*(l_i)] x_i l_i n_i \right)$$

$$\| x_{k+1} - x^* \|_A^2 = \sum_{i=l}^{n} l_i [1 + l_i P_k^*(l_i)]^2 x_i^2 \qquad \text{Orthogonal eigenvectors}$$

$$\| x_{k+1} - x^* \|_A^2 = \sum_{i=l}^{n} l_i [1 + l_i P_k^*(l_i)]^2 x_i^{\,2}$$

Since polynomial generated by GC is optimal

$$\| x_{k+1} - x^* \|_A^2 = \min_{P_k} \sum_{i=1}^{n} l_i [1 + l_i P_k(l_i)]^2 x_j^{\,2}$$

$$\| x_{k+1} - x^* \|_A^2 \le \min_{P_k} \max_{1 \le i \le n} [1 + l_i P_k(l_i)]^2 \left(\sum_{j=1}^{n} l_j x_j^{\,2} \right)$$

(C) $$\| x_{k+1} - x^* \|_A^2 \le \min_{P_k} \max_{1 \le i \le n} [1 + l_i P_k(l_i)]^2 \, \| x_0 - x^* \|_A^2$$

$$x_0 - x^* = \sum_{i=l}^{n} x_i v_i$$

(B) $$\min_{P_k} \max_{1 \le i \le n} [1 + l_i P_k(l_i)]^2$$

Convergence

$$x_0 - x^* = \sum_{i=1}^{n} x_i v_i$$

$$\| x_0 - x^* \|_A^2 = \sum_{i=1}^{n} x_i v_i^T A \sum_{i=1}^{n} x_i v_i$$

$$= \sum_{i=1}^{n} x_i v_i^T \sum_{i=1}^{n} x_i A v_i$$

$$= \sum_{i=1}^{n} x_i v_i^T \sum_{i=1}^{n} x_i l_i v_i$$

$$= \sum_{i=1}^{n} x_i^{\,2} l_i$$