MAT 228A Notes

Sam Fleischer

November 29, 2016

1 Conjugate Gradient Method

$Au = f$ where A is symmetric and positive definite (s.p.d.). It turns out u minimizes ϕ where

$$\phi(u) = u^T Au - u^T f.$$

Remember the definition

$$\kappa = \|A\|_2 \left\|A^{-1}\right\|_2.$$

For symmetric matrices,

$$\kappa = \frac{\max\limits_j |\lambda_j|}{\min\limits_j |\lambda_j|}.$$

In conjugate gradient method, we will not always go in the direction of steepest descent (the residual direction). Let p_k be the vector to follow.

$$u_{k+1} = u_k + \alpha p_k.$$

Follow this search direction until ϕ increases, i.e. minimize ϕ on a 1D space. We find

$$\alpha = \frac{p_k^T r_k}{p_k^T A p_k}.$$

1.1 Start in 2D

- Initial guess u_0
- compute residual r_0
- Initialize $p_0 = r_0$
- Get $u_1 = u_0 + \alpha p_0$
- Pick p_1 so that $p_1^T A p_0 = 0$.
  - In otherwords, p_0 and p_1 are A-conjugate, or

    $$p_1^T A p_0 = \langle p_0, p_1 \rangle_A$$

    because A is s.p.d.
  - Why do this? Notice that p_0 is tangent to the level set of ϕ at $\phi(u_1)$, i.e. we know $p_0^T r_1 = 0$. So,

    $$\begin{aligned} p_0^T r_1 &= 0 \\ p^T(f - Au_1) &= 0 \\ p^T(Au - Au_1) &= 0 \\ p^T A(u - u_1) &= 0 \end{aligned}$$

    So p_1 will be pointing in the direction of $u - u_1$. So u_2 is the exact solution.

1.2 3D

- initial guess u_0
- compute residual r_0
- Initialize $p_0 = r_0$
- Get $u_1 = u_0 + \alpha p_0$
- Pick p_1 to be A-conjugate to p_0.
  - This is a 2D space to choose from.. just pick a direction in that space.
- p_1 and r_1 define α
- calculate $u_2 = u_1 + \alpha p_1$.
  - p_0 and p_1 span a plane. This plane, which is $c_0 p_0 + c_1 p_1 + u_2$, is tangent to the level surface at $\phi(u_2)$.
  - At this point, we've minimized on this plane.
- Pick p_2 to be A-conjuage to $\{p_0, p_1\}$.
- Minimizing ϕ along this direction gives us the exact solution.

$$\begin{aligned}(c_0 p_0 + c_1 p_1)^T r_2 &= 0 \\ (c_0 p_0 + c_1 p_1)^T (f - Au_2) &= 0 \\ (c_0 p_0 + c_1 p_1)^T (Au - Au_2) &= 0 \\ (c_0 p_0 + c_1 p_1)^T A(u - u_2) &= 0\end{aligned}$$

So p_2 is pointing right at the solution.

1.3 In general

So CG will (no round off) give the xact solution in N steps. Usually we get close in far fewer steps.

1.3.1 Pseudocode

- Initialize u_0, $r_0 = f - Au_0$, $p_0 = r_0$
- loop in k
  - $w = Ap_k$
  - $\alpha = \dfrac{r_k^T r_k}{w^T p_k}$
  - $u_{k+1} = u_k + \alpha p_k$
  - $r_{k+1} = r_k - \alpha w$
  - check $\|r_{k+1}\|$ for stopping
  - compute $\beta = \dfrac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$
  - $p_{k+1} = r_{k+1} + \beta p_k$
    - This is r_{k+1} without components in $\{p_0, p_1, \ldots, p_{k-1}\}$.

2 Analysis of Convergence

Because A is s.p.d. it defines a norm $\|u\|_A^2 = u^T Au$. We can show $\|e_j\|_A \le 2\left(\dfrac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^j \|e_0\|_A$, which shows convergence since $\|e_j\|_A \to 0$ as $j \to \infty$. So, $\kappa \approx 1$ gives fast convergence and $\kappa \gg 1$ gives slow convergence.

2.1 Theorem from Book

The vectors in the CG algorithm have the following properties provided $r_k \neq 0$:

1. p_k is A-conjugate to all previous search directions.
2. r_k is orthogonal to all previous residuals.
3. The following subspaces of $\mathbb{R}^n$ are identical
   (a) span $\{p_0, p_1, \ldots, p_{k-1}\}$
   (b) span $\left\{r_0, Ar_0, A^2 r_0, \ldots, A_{k-1} r_0\right\}$
   (c) span $\left\{e_0, Ae_0, A^2 e_0, \ldots, A_{k-1} e_0\right\}$

   Define $K_n = \text{span}\left\{r_0, Ar_0, A^2 r_0, \ldots, A_{n-1} r_0\right\}$ as the n-dimensional Krylov space associated with r_0.

   $$u_n \in u_0 + K_n.$$

   u_n minimumzes ϕ in this space. Minimizing $\phi(u)$ is equivalent to minimizing $\|e_j\|_A^2$.

   $$\begin{aligned}
   \|e_j\|_A^2 &= e_j^T A e_j \\
   &= (u_j - u)^T A (u_j - u) \\
   &= u_j^T A u_j - 2u_j^T A u + u^T A u \\
   &= u_j^T A u - 2u_j^T f + u^T A u \\
   &= 2\phi(u_j) + C
   \end{aligned}$$