Discussion #5

Today:

1) Newton's Method
2) Conjugate Gradient

- HW4 Due Friday
- Midterm - Next week
  24 Hours

1) Newton's Method

Recall in 1D:

$$x_{k+1} = x_k - \boxed{\frac{f'(x_k)}{f''(x_k)}}$$

- Quadratic Convergence
- Uses First Derivative and Second Derivative Information
- Convergence is slow

$\boxed{1.\ f''(x^*) = 0}$

2. f'''

$f(x) = x^{4/3}$

2. $|f''|$ is large, $f(x) - x$

3. $|x_k - x^*|$ is large

Error Term from Taylor expansion

is $\frac{1}{2} \frac{f'''(\xi)}{f''(x^*)} \cdot (x_k - x^*)^2$

- n dimensions

$$x_{k+1} = x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k)$$

Convergence Issues may happen in all of the above cases.

In particular, want $\nabla^2 f(x_k)$ invertible, preferably want it to be positive definite.

Order of Convergence:

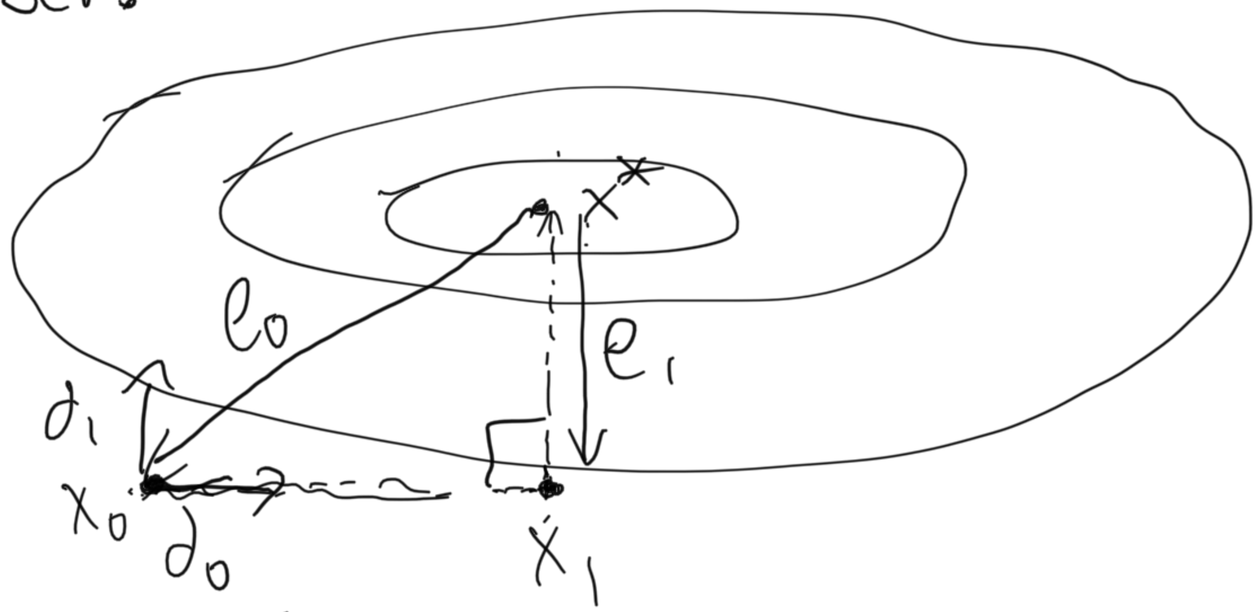
$$0 < \lim_{k \to \infty} \frac{|x_k - x^*|}{|x_k - x^*|^p} < \infty$$

Conjugate Gradient Method

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

Choose $d_0, d_1, \ldots, d_{n-1}$ orthogonal directions,

At each step, choose step size to get as close as possible to the minimizer.



Want $d_0^T e_1 = 0$

$$e_1 = e_0 + \alpha_0 d_0$$

$$d_0^T (e_0 + \alpha_0 d_0) = 0$$

$$\Rightarrow \boxed{\alpha_0 = \frac{\cdots}{d_0^T d_0}}$$

Doesn't work because need to know e_0.

We do know $\underline{Qe_0}$

$$e_0 = x_0 - x^* = x_0 - Q^{-1}b$$

$$\Rightarrow \quad Qe_0 = Qx_0 - b$$

$$= \nabla f(x_0)$$

$$= g_0$$

Instead, chose $\underline{d_0, \ldots, d_{n-1}}$ to be Q-orthogonal.

Require $d_0^T Q e_1 = 0$

$$\Rightarrow \underline{\alpha_0} = \frac{-d_0^T Q e_0}{d_0^T Q d_0}$$

$$= \cdots$$

$$= \frac{d_0^T g_0}{d_0^T Q d_0}$$

We can decompose

$$e_0 = \boxed{\underline{a_0} d_0} + \dots + a_{n-1} d_{n-1}$$

$$d_0^T Q e_0 = a_0 d_0^T Q d_0$$

$$\Rightarrow a_0 = \frac{d_0^T g_0}{d_0^T Q d_0} = -\alpha_0$$

$$\Rightarrow e_1 = a_1 d_1 + \dots + a_{n-1} d_{n-1}$$

$$e_2 = a_2 d_2 + \dots + a_{n-1} d_{n-1}$$

$$e_i = \sum_{j=i}^{n-1} a_j d_j$$

How to find $d_0, \dots, d_{n-1}$?

Suppose we have

$$\underline{u_0, u_1, \dots, u_{n-1}}$$

Use Gram Schmidt.

Set $d_0 = u_0$

For $i = 1, \dots, n-1$

Set $d_i = u_i - \sum_{j=1}^{i-1} \frac{\langle u_i, d_j \rangle}{\langle d_j, d_j \rangle} d_j$

To make d_i orthogonal,

use $\langle u, v \rangle = u^T v$

To make d_i Q-orthogonal

use $\langle u, v \rangle = u^T Q v$

$$d_i = u_i - \sum_{j=1}^{i-1} \frac{u_i^T Q d_j}{d_j^T Q d_j} d_j$$

Lemma: $g_i^T d_k = 0$ if $k < i$

Recall, $e_i = \boxed{\sum_{j=i}^{n-1} a_j d_j}$

Then $\underbrace{d_k^T Q e_i}_{\parallel} = 0$ for $k < i$.

$d_k^T g_i$

For CG, choose $\boxed{u_i = -g_i}$

Then $d_i = -g_i + \sum_{j=0}^{i-1} \frac{g_i^T Q d_j}{d_j^T Q d_j} d_j$

Then $\boxed{g_k^T d_i} = -g_k^T g_i + \sum_{j=0}^{i-1} \frac{g_i^T Q d_j}{d_j^T Q d_j} \boxed{g_k^T d_j}$

If $k > i$,

$\Rightarrow g_k^T g_i = 0$

Fact: $e_i = e_{i-1} + \alpha_{i-1} d_{i-1}$

Multiply by Q,

$$g_i = g_{i-1} + \alpha_{i-1} Q d_{i-1}$$

$\Rightarrow g_k^T g_i = g_k^T g_{i-1} + \alpha_{i-1} g_k^T Q d_i$

If $k \neq i, i-1$,

$$\Rightarrow \alpha_{i-1} g_k^T Q d_i = 0$$

$\Rightarrow$ g_k is Q-orthogonal to d_i if $k \neq i, i-1$

$$d_i = -g_i + \sum_{j=0}^{i-1} \frac{g_i^T Q d_j}{d_j^T Q d_j} d_j$$

$$= -g_i + \frac{g_i^T Q d_{i-1}}{d_{i-1}^T Q d_{i-1}} d_{i-1}$$