Newton's Method

~~Rec~~ Recall in 1D:

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

~~Issues~~:

Issues when

$$f''(x_k) = 0$$

In n dimensions

$$d_k = -\nabla^2 f(x_k)^{-1} \nabla f(x_k)$$

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

Issue if $\nabla^2 f(x_k)$ is not invertible or not positive definite

If $\nabla^2 f(x_k) > 0$, $f(x_{k+1}) < f(x_k)$
$\nabla f(x_k) \neq 0$

~~$f(x_{k+1})$~~ ~~$\phi(x_k)$~~ $\phi(\alpha) = f(x_k + \alpha d_k)$

$$\phi'(\alpha) = \nabla f(x_k + \alpha d_k)^T d_k$$

$$\phi'(0) = -\nabla f(x_k)^T \left(\nabla^2 f(x_k)\right)^{-1} \nabla f(x_k) < 0$$
$$= \nabla f(x_k)^T d_k$$

Hence d_k is a descent direction

We can modify algorithm

$$x_{k+1} = x_k - \underbrace{\left(\nabla^2 f(x_k) + \lambda I\right)}^{-1} \nabla f(x_k)$$

Take λ large enough so this is positive def.

Conjugate Gradient

$f(x) = \frac{1}{2}x^TQx - b^Tx$

$\nabla f(x) = Qx - b$

Solution x^* satisfies

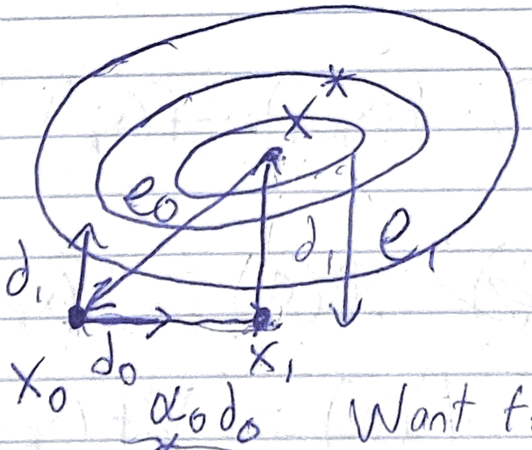
$Qx^* = b$

$= Qx - Qx^*$

$= Q(x - x^*)$

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

At each step choose step size α to get close as possible to minimizer



$e_0 = x_0 - x^*$

$e_1 = x_1 - x^*$

Want to choose α_0 so that error after 1 step is orthogonal to d_0

Want $d_0^T e_1 = 0$

$e_1 = e_0 + \alpha_0 d_0$

$d_0^T e_1 = d_0^T e_0 + \alpha_0 d_0^T d_0 = 0$

$\Rightarrow \quad \alpha_0 = \dfrac{-d_0^T e_0}{d_0^T d_0}$

Problem, don't know e_0.

We do know $Qe_0 = Q(x_0 - x^*)$
$= \nabla f(x_0)$

Instead, choose $d_0, \ldots, d_{n-1}$ to be Q-orthogonal
i.e, $d_i^T Q d_j = 0$

Instead need $d_0^T Q e_1 = 0$

$$\begin{aligned} &\parallel \\ &d_0^T Q(e_0 + a_0 d_0) \\ \Rightarrow\ & d_0^T \underbrace{Q e_0}_{g_0} + \alpha_0 d_0^T Q d_0 \end{aligned}$$

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

Since $d_0, \ldots, d_n$ linearly independent

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

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

$$\Rightarrow a_0 = \frac{d_0^T Q e_0}{d_0^T Q d_0}$$

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

Gram Schmidt

Take $u_0, \ldots, u_{n-1}$ linearly independent

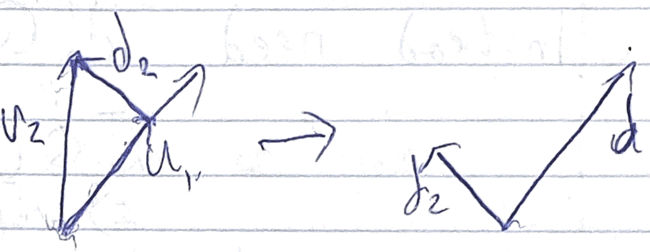
Set $d_0 = u_0$

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

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



To make d_i Q-orthogonal

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

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

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

Recall $e_i = \sum_{j=i}^{n-1} a_j d_j = a_i d_i + \dots + a_{n-1} d_{n-1}$

$d_k^T Q e_i = 0$ for $k < i$

$\quad \|$

$d_k^T g_i$