Conjugate Gradient Method I

- For symmetric positive definite matrices only
- One of the most frequently used iterative methods
- Before introducing CG, we discuss a related method called the steepest descent method
- We still consider solving

$$\min_x \frac{1}{2}x^T Ax - b^T x$$

Steepest Descent Method I

- Gradient direction

$$
\begin{aligned}
f(x) =& f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2 + \cdots \\
f(x) =& f(x_k) + \nabla f(x_k)^T (x - x_k) + \\
& \frac{1}{2}(x - x_k)^T \nabla^2 f(x_k)(x - x_k) + \cdots
\end{aligned}
$$

Omit $\frac{1}{2}(x - x_k)^T \nabla^2 f(x_k)(x - x_k) + \cdots$

$$
f(x) \approx f(x_k) + \nabla f(x_k)^T (x - x_k)
$$

Steepest Descent Method II

- Minimize

  $$\nabla f(x_k)^T (x - x_k)$$

  If

  $$\min_{\|x - x_k\| = 1} \nabla f(x_k)^T (x - x_k),$$

  then

  $$x - x_k = \frac{-\nabla f(x_k)}{\|\nabla f(x_k)\|}$$

  is the direction.

Steepest Descent Method III

- We need
$$\|x - x_k\| = 1$$
Otherwise the minimization goes to $-\infty$
- Now
$$\nabla f(x_k) = Ax_k - b.$$
- Let
$$r = -\nabla f(x_k) = b - Ax_k, x = x_k$$

Steepest Descent Method IV

- Minimize along the direction

$$\min_{\alpha} \frac{1}{2}(x+\alpha r)^T A(x+\alpha r) - (x+\alpha r)^T b$$

$$\min_{\alpha} \frac{1}{2}\alpha^2 r^T Ar + \alpha r^T Ax - \alpha r^T b$$

$$\min_{\alpha} \frac{1}{2}\alpha^2 r^T Ar + \alpha r^T (Ax - b)$$

Steepest Descent Method V

- A problem of one variable:

$$\alpha r^T A r + r^T (Ax - b) = 0$$
$$\alpha = \frac{r^T (b - Ax)}{r^T A r}$$

- Note

$$r^T A r \neq 0 \text{ if } A \text{ is positive definite and } r \neq 0$$

Steepest Descent Method VI

- Now $r = b - Ax_k$

$$\begin{aligned}\alpha =& \frac{(b - Ax_k)^T(b - Ax_k)}{(b - Ax_k)^T A(b - Ax_k)}\\ =& \frac{r^T r}{r^T A r}\end{aligned}$$

- The algorithm:

Steepest Descent Method VII

$$
\begin{aligned}
&k = 0;\ x_0 = 0;\ r_0 = b \\
&\text{while } r_k \neq 0 \\
&\qquad k = k + 1 \\
&\qquad \alpha_k = r_{k-1}^T r_{k-1} / r_{k-1}^T A r_{k-1} \\
&\qquad x_k = x_{k-1} + \alpha_k r_{k-1} \\
&\qquad r_k = b - A x_k \\
&\text{end}
\end{aligned}
$$

- It converges but may be very slow

General Search Directions I

- Suppose we have x_k obtained by

$$\min_{\alpha} f(x_{k-1} + \alpha p_k),$$

where p_k is any direction

$$\begin{aligned}
& f(x_{k-1} + \alpha p_k) \\
= \ & \frac{1}{2}(x_{k-1} + \alpha p_k)^T A(x_{k-1} + \alpha p_k) - b^T(x_{k-1} + \alpha p_k) \\
= \ & \text{constant } + \alpha p_k^T(Ax_{k-1} - b) + \frac{1}{2}\alpha^2 p_k^T A p_k
\end{aligned}$$

General Search Directions II

$$\alpha = \frac{p_k^T(r_{k-1})}{p_k^T A p_k}$$

- A more general algorithm:

$k = 0$; $x_0 = 0$; $r_0 = b$
while $r_k \neq 0$
  $k = k + 1$
  Choose a direction p_k such that $p_k^T r_{k-1} \neq 0$
  $\alpha_k = p_k^T r_{k-1} / p_k^T A p_k$
  $x_k = x_{k-1} + \alpha_k p_k$
  $r_k = b - A x_k$
end

General Search Directions III

- By this setting

$$x_k \in x_0 + \text{span}\{p_1, \ldots, p_k\}$$

- The question is then how to choose suitable directions?