1.3 Successive Over-Relaxation (SOR)

$$\boxed{x^{(k+1)} = D^{-1}\left(b - Lx^{(k+1)} - Ux^{(k)}\right)} \leftarrow \text{GS}$$

$$\cancel{x^{(k+1)} = (D+L)^{-1}\left(b - Ux^{(k)}\right)} \leftarrow \text{~~another form of GS~~}$$

$$\boxed{x^{(k+1)} = x^{(k)} + \omega \underbrace{D^{-1}\left(b - Lx^{(k+1)} - Ux^{(k)} - Dx^{(k)}\right)}_{\text{correction term leading } x^{(k)} \to x^{(k+1)}}}$$

SOR : $1 < \omega < 2$

convergent for $0 < \omega < 2$

Rmk

i) $\omega = 1$; GS

ii) $0 < \omega < 1$: interpolation ("under"-relaxation)

iii) $\boxed{1 < \omega < 2}$: extrapolation ("over"-relaxation)

iv) convergent for $0 < \omega < 2$ if A is p.d.

v) SOR diverges if $\omega \geq 2$ or $\omega \leq 0$

vi) If A is p.d & tridiagonal,

the optimal $\omega_{opt} = \dfrac{2}{\left(1+\sqrt{1-\rho(D^{-1}R)^2}\right)}$

where $A = L + D + U$,

$R = A - D = L + U$

Rmk. GJ parallel vs. SOR serial

in terms of

convergence rate vs scalability.

$\boxed{2}$ Conjugate Gradient method (CG)

only for spd.

(i) $O(m^3)$ in the worst case

(same as the direct methods)

(ii) $O(Nm^2)$ in practice, $N \ll m$

§2.1 Real s.p.d linear systems & minimization

Thm. Consider $Ax=b$, A: real, spd.

Then the soln to $Ax=b$ is obtained by solving the minimization prob. of a quad. form $f(x) = \frac{1}{2}x^TAx - x^Tb$.

pf. $f: \mathbb{R}^m \longrightarrow \mathbb{R}$,

if A is spd, then f: convex

$\rightarrow$ $\exists$ single global min, namely x^*.

$\rightarrow$ $0 = \nabla f = Ax - b$ ∎

Ex. $m=2$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$; $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$

A: spd

$$f(x) = \frac{1}{2}x^TAx - x^Tb$$

$$= \frac{1}{2}\left(a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2\right) - (x_1b_1 + x_2b_2)$$

$$\rightarrow \quad 0 = \left.\frac{\partial f}{\partial x_1}\right|_{x^*} = a_{11}x_1^* + a_{12}x_2^* - b_1$$

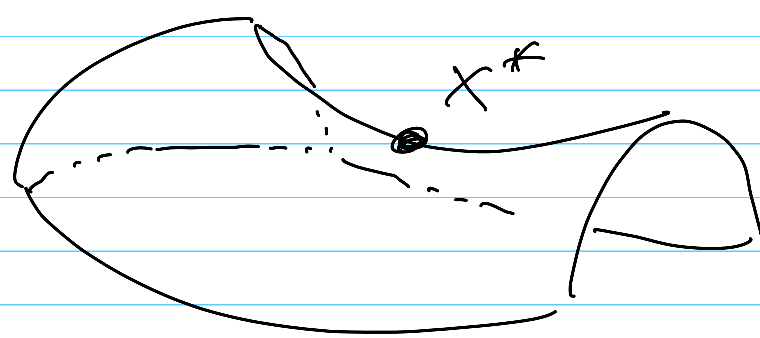
$$0 = \left.\frac{\partial f}{\partial x_2}\right|_{x^*} = a_{22} x_2^* + a_{12} x_1^* - b_2$$

$$\Leftrightarrow \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

<u>Rmk.</u> If A is <u>indefinite</u> (neither p.d. or n.d)

$\rightarrow \exists\, x^*$ s.t $\nabla f(x^*) = 0$ but

x^* is a <u>saddle</u> pt.



$\rightarrow$ We don't consider this case!

<u>§2.2.</u> Iteration methods for min. quad. ftn f,

$\rightarrow$ Consider g s.t $x^{(k+1)} = g(x^{(k)})$, which gradually decrease $\|x^{(k+1)} - x^*\|$

→ Since quad ftn f is convex,

i) begin with a particular "search direction"

ii) search distance
along the search direction
to get to a local min

iii) repeat (i) & (ii) until getting
to a global min.

→ let $p^{(k)}$: search direction

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$$

α_k → search distance

$$\rightarrow f(x^{(k+1)}) = \frac{1}{2}(x^{(k)} + \alpha_k p^{(k)})^T A (x^{(k)} + \alpha_k p^{(k)}) - (x^{(k)} + \alpha_k p^{(k)})^T b$$

→ droping " k " index,

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

$$+ \frac{\alpha}{2} \left(p^T A x + x^T A p \right)$$

$\alpha (p^T A x)$

$$+ \frac{\alpha^2}{2} \left(p^T A p - \alpha p^T b \right)$$

$$\rightarrow 0 = \frac{d f(x^{(k+1)})}{d \alpha_k} \quad \text{then}$$

$$\boxed{p^T A x + \alpha p^T A p - p^T b = 0}$$

$$\rightarrow \alpha_k = \frac{p^{(k)T} r^{(k)}}{p^{(k)T} A p^{(k)}}, \quad r^{(k)} = b - A x^{(k)}$$

Now we conside to find $p^{(k)}$:

① $p^{(k)} = \underbrace{-\nabla f(x^{(k)})}_{= r^{(k)}}$; steepest gradient descent

② $p^{(k)} = r^{(k)} + \beta_{k-1} p^{(k-1)}$; A- conjugate direction

(CG)

§2.3. Steepest gradient descent

$$p^{(k)} = -\nabla f(x^{(k)}) = -(Ax^{(k)} - b) = r^{(k)}$$

Alg. Initialize $x = 0$

$r = b$

do while $\|r\| >$ large

$p = r$

$$\alpha = \frac{p^T r}{p^T A p}$$

$x = x + \alpha p$

$r = b - Ax$

enddo

§2.4 The (A-Conjugate Gradient) Method

(CG)

Def. Two non-zero vectors u & v :

A-conjugate if

$$(u, Av) = u^T A v = 0$$

Rmk. If A is p.d, we can define a new norm:

$$\|u\|_A = (u, Au)^{\frac{1}{2}} = \sqrt{u^T A u}$$

Def. A set of vectors $\{P_i\}$: a conj. set (w.r.t. A) if

$$(P_i, AP_j) = P_i^T A P_j = 0, \quad \forall i \neq j.$$

Rmk. A: real sym. $m \times m$.

$\rightarrow$ $\begin{cases} \lambda_i : \text{real (eig vals)} \\ v_i : \text{orthogonal eig vectors.} \end{cases}$

$\rightarrow$ $(v_i, Av_j) = v_i^T A v_j$

$$= v_i^T \lambda_j v_j$$

$$= \lambda_j v_i^T v_j = 0, \quad \forall i \neq j$$

→ In general, any real symm. matrix has a conj. set $\{p_i\}_{i=0}^{m-1}$ and they form a basis for $\mathbb{R}^m$.

<u>Rmk</u>. (Prob 8, HW2)

Supp. a set $\{p_i\}_{i=0}^{m-1}$: a conj. set for A ← spd.

$$\rightarrow \quad x = \sum_{i=0}^{m-1} \alpha_i p_i$$

soln to $Ax = b$

$$\rightarrow \quad (p_j, Ax) = (p_j, b) = p_j^T b$$

$$\| $$

$$\left(p_j, A \sum \alpha_i p_i\right) = p_j^T A \sum \alpha_i p_i$$

$$= p_j^T \sum \alpha_i A p_i$$

$$= \alpha_j \, p_j^T A p_j$$

$$\rightarrow \alpha_j = \frac{p_j^T b}{p_j^T A p_j}$$

$$\rightarrow x = \sum_{i=0}^{n-1} \left(\frac{p_i^T b}{p_i^T A p_i} \right) p_i$$

Not good ☹ expensive

Rmk. We hope that $\exists$ iterative method that gives us $p^{(0)}, p^{(1)} \dots$ iteratively.

Let
$$\begin{cases} p^{(0)} = p_0 & \text{at } k=0 \\ p^{(1)} = p_1, & \text{at } k=1, \text{ etc.} \end{cases}$$

Let
$$x^{(k+1)} \stackrel{def}{\equiv} \sum_{i=0}^{k} \alpha_i \, p^{(i)}$$

$$\stackrel{def}{\equiv} x^{(k)} + \alpha_k \, p^{(k)} \quad \text{: successive appn}$$

$\rightarrow$ $x^{(k+1)} \rightarrow x$: soln to $Ax = b$

when $k+1 = m$.

$\alpha_k \rightarrow 0$ in the limit

$\hookrightarrow$ this happens in N steps

then $O(N m^2)$, $N \ll m$.

(i) $p^{(k)} \neq -\nabla f(x^{(k)}) \rightarrow$ steepest grad. descent

(ii) $p^{(k)}$: A - conj. to each other

i.e, $p^{(k)T} A p^{(i)} = 0, \ \forall i = 0, \dots k-1$

For $k=0$; $x^{(0)} = 0$

$p^{(0)} = r^{(0)} = b - Ax^{(0)} = b$

same as in the steepest grad. descent.

For $k=1$;

$$x^{(1)} = x^{(0)} + \alpha_0 p^{(0)}$$

$$\boxed{\underline{p}^{(1)} = \underline{r}^{(1)} = b - Ax^{(1)}}$$ (Steepest descent)

$$p^{(1)T} A p^{(0)} = \underline{r}^{(1)T} A \underline{b}$$

$$= (b - Ax^{(1)})^T A b$$

$$= b^T A b - x^{(1)T} A^T A b$$

is not $\neq 0$. in general

A-conj.

$$p^{(1)} = r^{(1)} + \beta_0 p^{(0)}$$

We determine β_0 s.t $p^{(1)T} A p^{(0)} = 0$

$$0 = (r^{(1)} + \beta_0 p^{(0)})^T A p^{(0)}$$

$$= r^{(1)T} A p^{(0)} + \beta_0 p^{(0)T} A p^{(0)}$$

$$\therefore \quad \beta_0 = - \frac{r^{(1)T} A p^{(0)}}{p^{(0)T} A p^{(0)}}$$

<u>Alg</u>. The CG alg.

$$x^{(0)} = 0$$

$$r^{(0)} = p^{(0)} = b$$

$\Rightarrow$ (i) $$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}$$

$$\alpha_k = \frac{p^{(k)T} r^{(k)}}{p^{(k)T} A p^{(k)}}$$

(ii) $$r^{(k+1)} = b - A x^{(k+1)}$$

(iii) $$p^{(k+1)} = r^{(k+1)} + \beta_k p^{(k)}$$

$$\beta_k = - \frac{r^{(k+1)T} A p^{(k)}}{p^{(k)T} A p^{(k)}}$$

$\longrightarrow$ this converges to the true soln in at most m iterations