Newton Methods

Note Title

$$\left[\; p_k = -B_k^{-1} \nabla f_k \;\right]$$

NM $\quad B_k = \nabla^2 f_k$

QN $\quad B_k \approx \nabla^2 f_k$

Zoutendijk's cond$\underline{n}$

$$f_{k+1} \le f_0 - c \underbrace{\sum_{j=0}^{k} \cos^2\theta_j \, \|\nabla f_j\|^2}.$$

15 - Windows Journal

File Edit View Insert Actions Tools Help

Page Width

$$\lim_{j\to\infty} \underbrace{\cos^2\theta_j \, \|\nabla f_j\|^2}_{} \to 0 \qquad \text{for bounded fns}$$

either $\cos\theta_j \to 0$ or $\boxed{\|\nabla f_j\| \to 0}$

we want

(stationary pt is $\nabla f = 0$)

$$\cos^2\theta_j > 0 \quad \Leftarrow$$

imp

Condition on the Hessian for this ( ) to happen.

15 - Windows Journal

File Edit View Insert Actions Tools Help

$$K(B_k) \leqslant M \qquad \text{Cond}\underline{n} \text{ no.}$$

Proof:

$$\cos\theta_k = \frac{-\nabla f_k^T p_k}{\|\nabla f_k\| \, \|p_k\|}$$

2 properties.

1) If B exists & is PD

$$\sqrt{B} = B^{1/2}.$$

Proof:

$$\cos\theta_k = \frac{-\nabla f_k^T p_k}{\|\nabla f_k\| \, \|p_k\|}$$

2 properties

1) If B exists & is P.D.

$$\sqrt{B} = B^{1/2} = Q\Lambda^{1/2}Q^T$$

$Q\Lambda Q^T$ is also P.D.

$B^{-1/2}$ is also P.D.

(dropping k)

② Norms $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

$$p = -B^{-1}\nabla f \Rightarrow \nabla f = -Bp.$$

③ $K(B) = \|B\| \, \|B^{-1}\| \leq M$ given

15 - Windows Journal

File Edit View Insert Actions Tools Help

$$\cos\theta = \frac{p^T B p}{\|Bp\|\,\|p\|} \geq \frac{p^T B p}{\|B\|\,\|p\|^2} \qquad \frac{1}{\|A\|\,\|B\|} \leq \frac{1}{\|A \cdot B\|}$$

$$p^T B p = p^T \sqrt{B}\sqrt{B}\, p = \|\sqrt{B}\, p\|^2$$

$$\cos\theta \geq \frac{\|B^{1/2} p\|^2}{\|B\|\,\|p\|^2}$$

$$B^{-1/2} \cdot B^{1/2} \cdot p$$

15 - Windows Journal

File Edit View Insert Actions Tools Help

$$\cos\theta \geq \frac{\|B^{1/2}p\|^2}{\|B\|\,\|p\|^2} \geq \frac{\|B^{1/2}\cancel{p}\|^2}{\|B\|\left[\|B^{-1/2}\|\,\|B^{1/2}\cancel{p}\|\right]^2}$$

$$\left[\underbrace{B^{-1/2}}_{a}\,\underbrace{B^{1/2}p}_{b}\right]$$

$$\|ab\| \leq \|a\|\,\|b\|$$

$$\|B^{-1/2}\|^2 = \|B^{-1}\|$$

$$\geq \frac{1}{\|B\|\,\|B^{-1}\|} \geq \frac{1}{M}$$

$$\Rightarrow \|\nabla f\|^2 \to 0 \quad \text{as } j \to \infty$$

i.e. NM converges.

$$\Rightarrow \|\nabla f\|^2 \to 0 \quad \text{as } j \to \infty$$

i.e. NM converges

Rate of convg is quadratic → NW Sec 3.5

quad

$$\|x_{k+1} - x^*\|^2 \le \left(\qquad \right) \|x_k - x^*\|^2$$

Summarize NM:

1) Pick x_0

2) $p_k = -B_k^{-1} \nabla f_k$ dir.

Check B_k for P.D, then

3) Compute $\alpha_k \longrightarrow$ B.T. L.S.

4) Go to x_{k+1}. 5) Check $\|\nabla f_k\|$ $\nearrow$ END $\searrow$ STEP 2.

If B_k is NOT P.D. then what to do?

1. Drop down into S.D. or CG or NLCG.
2. Hessian modification.

↳ How to check if B_k is P.D

a) eig values $\quad O(n^3) \quad$ $x^T B_k x > 0, \forall x$

b)

15 - Windows Journal

File Edit View Insert Actions Tools Help

Page Width

How to check if B_k is P.D

a) eig values $O(n^3)$ $\left(x^T B_k x > 0, \forall x\right)$

b) Cholesky decomp.

If A is P.D, then $A = LL^T$ — L lower tri, diags are > 0

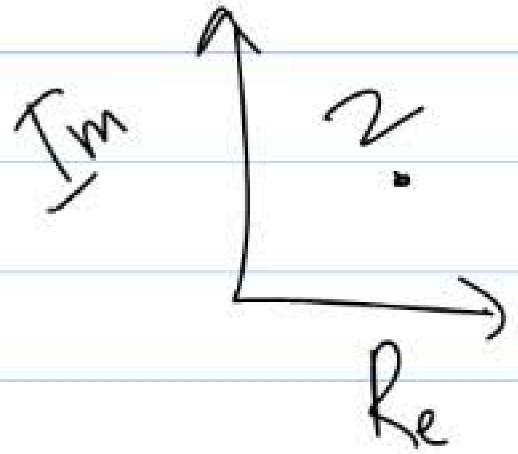
If $A = LL^T$, then A is P.D.

c)

If $A = LL^T$, then A is P.D.

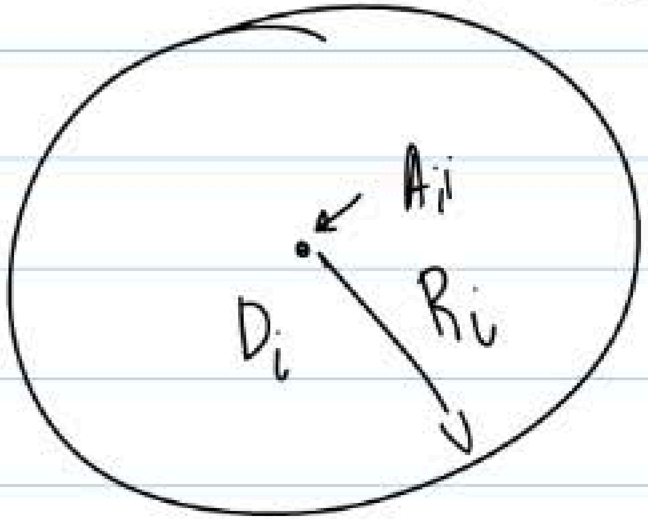
c) Gershgorin disk thm (L.A.)

Assume A is a complex matrix.



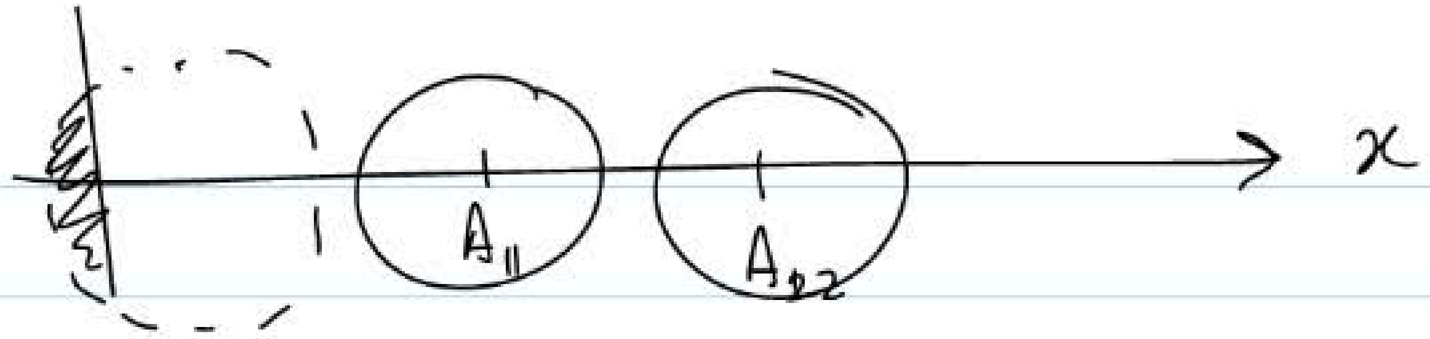
Disk$_i$ $\rightarrow$ Centre A_{ii}

Radius $R_i = \sum_{\substack{i \neq j \\ j=1}}^{n} |A_{ij}|$



Thm: Every eig value of A lies within atleast one disc D_i.

B_k is real valued.



$\Rightarrow$ Easy to check if $\lambda_{min} > 0$

Hessian Modification

$A \rightarrow$ is not P.D.

perturb $A + \Delta A$, s.t. it becomes P.D.

Require $\|\Delta A\|$ be as low as possible

15 - Windows Journal

File Edit View Insert Actions Tools Help

$$A = Q \Lambda Q^T \qquad \Delta A = \underbrace{\tau}_{\text{Scalar}} I = \tau Q Q^T$$

$$A + \Delta A = Q \underbrace{[\Lambda + \tau I]} Q^T$$

$$\lambda_{min} + \tau \geq \delta$$

Saves you from bad Cond no.

$$\Rightarrow \tau \geq \delta - \lambda_{min}.$$

only if $\lambda_{min} < 0$.

$$\|\Delta A\| = \tau$$

Hessian modification :

$$A + \Delta A = Q\,[\underbrace{\Lambda + \tau I}]\,Q^T$$

Saves you from bad cond no.

$$\lambda_{min} + \tau \geqslant \delta$$

$$\Rightarrow \underline{\tau \geqslant \delta - \lambda_{min}}$$

only if $\lambda_{min} < 0$

$$\|\Delta A\|_2 = \tau$$

<u>Hessian modification</u>