Week 4, Lecture 2

§ 6.2 Application of Zoutendijk's Theorem to Quasi-Newton methods:

$$B_k p_k = -\nabla f_k \qquad (1)$$

Exam

Theorem 6.2 Consider an iterative method where the descent direction is given by (1), where B_k is a symmetric positive-definite matrix satisfying

$$\|B_k\|_2 \, \|B_k^{-1}\|_2 \leq M \qquad (2)$$

where M is a constant. Then

$$\cos\theta_k \geq 1/M .$$

<u>Remark</u>:

$$\kappa(B_k) = \|B_k\|_2 \, \|B_k^{-1}\|_2$$

is the <u>condition number</u> of the matrix B_k.

Here, $\|\cdot\|_2$ is the matrix norm (L^2 norm):

$$\|B_k\|_2 = \sup_{\|\underline{u}\|_2 = 1} \|B_k \underline{u}\|_2$$

When the matrix B_k is symmetric positive-definite, with positive eigenvalues $(\lambda_1, \ldots, \lambda_n)$, and $\max(\lambda_1, \ldots, \lambda_n) = \lambda_{max}$, then

$$\| B_k \|_2 = \lambda_{max} .$$

Similarly, if $\min(\lambda_1, \ldots, \lambda_n) = \lambda_{min}$, then

$$\| B_k^{-1} \|_2 = \frac{1}{\lambda_{min}} .$$

Then,

$$\kappa(B_k) = \| B_k \|_2 \| B_k^{-1} \|_2 = \frac{\lambda_{max}}{\lambda_{min}} \quad (3)$$

<u>Return to proof of theorem:</u>

$$\cos \theta_k = \frac{- \langle p_k, \nabla f_k \rangle}{\| \nabla f_k \|_2 \| p_k \|_2}$$

$$\stackrel{(i)}{=} \frac{+ \langle B_k^{-1} \nabla f_k, \nabla f_k \rangle}{\| \nabla f_k \|_2 \| B_k^{-1} \nabla f_k \|_2}$$

Expand in terms of eigenbasis, as B_k is symmetric:

$$\nabla f_k = \sum_i x_i \underline{u}_i \, , \qquad x_i = \langle \underline{u}_i, \nabla f_k \rangle$$

$$B_k \underline{u}_i = \lambda_i \underline{u}_i$$

All eigenvalues are strictly positive:

$$B_k^{-1} \underline{u}_i = \frac{1}{\lambda_i} \underline{u}_i$$

$$\Rightarrow B_k^{-1} \nabla f_k = \sum_i x_i \cdot \frac{1}{\lambda_i} \underline{u}_i$$

Sub back into expression for $\cos \theta_k$:

$$\cos \theta_k = \frac{\langle B_k^{-1} \nabla f_k, \nabla f_k \rangle}{\|\nabla f_k\|_2 \, \|B_k^{-1} \nabla f_k\|_2}$$

$$= \frac{\left\langle \sum_i \frac{x_i}{\lambda_i} \underline{u}_i , \sum_j x_j \underline{u}_j \right\rangle}{\left\|\sum x_i \underline{u}_i\right\|_2 \left\| \sum_i \frac{x_i}{\lambda_i} \underline{u}_i \right\|_2}$$

Orthonormal eigenbasis:

$$\langle \underline{u}_i , \underline{u}_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Hence,

$$\lambda_i \leq \lambda_{max}$$

$$\frac{1}{\lambda_i} \geq \frac{1}{\lambda_{max}}$$

$$\cos \theta_k = \frac{\sum_i \frac{x_i^2}{\lambda_i}}{\left(\sum_i x_i^2\right)^{1/2} \left(\sum_i \frac{x_i^2}{\lambda_i^2}\right)^{1/2}}$$

$$(\lambda_{min}^2)^{1/2}$$

$$\lambda_{min}$$

$$= |\lambda_{min}|$$

$$\geq \frac{\frac{1}{\lambda_{max}} \sum_i x_i^2}{\left(\sum_i x_i^2\right)^{1/2} \left(\sum_i x_i^2/\lambda_i^2\right)^{1/2}}$$

CamScanner

$$\cos\theta_k \geq \frac{\frac{1}{\lambda_{max}} \cancel{\sum_i x_i^2}}{\frac{1}{\lambda_{min}} \cancel{\left(\sum_i x_i^2\right)^{1/2}} \cancel{\left(\sum_i x_i^2\right)^{1/2}}}$$

$$\Longrightarrow \cos\theta_k \geq \frac{\lambda_{min}}{\lambda_{max}} = \frac{1}{\kappa(B_k)} \geq \frac{1}{M}.$$

Hence, $\boxed{\cos\theta_k \geq \frac{1}{M}}$ $\blacksquare$