LECTURE 7

NEWTON'S METHOD FOR MULTIVARIATE FUNCTIONS (Gill & Wright 2.4)

$F: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is smooth such that

$$F(x) = \begin{bmatrix} F_1(x) \\ F_2(x) \\ \vdots \\ F_m(x) \end{bmatrix}$$

AIM: Find a direction $p \in \mathbb{R}^n$ such that

$$F(x+p) \approx 0$$

(WILL REPLACE F WITH A LINEAR FUNCTION L TO FIND SUCH A p)

Let $\phi_j : \mathbb{R} \to \mathbb{R}$,

$$\phi_j(\alpha) = F_j(x+\alpha p) \qquad j=1,\dots,m.$$

By fundamental thm

$$\phi_j(1) = \phi_j(0) + \int_0^1 \phi_j'(\alpha)\, d\alpha$$

$$\Longrightarrow F_j(x+p) = F_j(x) + \int_0^1 \nabla F_j(x+\alpha p)^T p \, d\alpha$$

Consequently

$$F(x+p) = \begin{bmatrix} \overbrace{F_1(x) + \int_0^1 \nabla F_1(x+\alpha p)^T p \, d\alpha}^{F_1(x+p)} \\ \vdots \\ \underbrace{F_m(x) + \int_0^1 \nabla F_m(x+\alpha p)^T p \, d\alpha}_{F_m(x+p)} \end{bmatrix}$$

$$= \begin{bmatrix} F_1(x) \\ \vdots \\ F_m(x) \end{bmatrix} + \int_0^1 \underbrace{\begin{bmatrix} \nabla F_1(x+\alpha p)^T \\ \vdots \\ \nabla F_m(x+\alpha p)^T \end{bmatrix}}_{F'(x+\alpha p)} p \, d\alpha$$

<u>THM</u> (Taylor's thm with integral remainder)

Let $F: \mathbb{R}^n \to \mathbb{R}^m$ be smooth. Then

$$F(x+p) = F(x) + \int_0^1 F'(x+\alpha p)^T p \, d\alpha$$

Suppose p is small in norm. By continuity of $F'(x)$ we have

$$F'(x+\alpha p) \approx F'(x) \quad \text{for all } \alpha \in [0,$$

It follows from Taylor's thm that

$$F(x+p) = F(x) + \int_0^1 F'(x+\alpha p)p\,d\alpha$$

$$\approx F(x) + \int_0^1 F'(x)p\,d\alpha$$

$$= F(x) + F'(x)p$$

Since we aim for a $p \in \mathbb{R}^n$ such that $F(x+p) \approx 0$, direction p satisfies

$$F'(x)p \approx -F(x)$$

NEWTON ITERATION

(1) Solve

$$F'(x_k)p_k = -F(x_k) \quad \left(\begin{matrix} m\times n \text{ linear} \\ \text{system} \end{matrix}\right)$$

for p_k (called Newton direction)

(2) $x_{k+1} = x_k + p_k$

EXAMPLE

$f: \mathbb{R}^3 \to \mathbb{R}$, $f(x) = 2x_1^2 - x_2^2 + 2x_3^2 + 4x_1x_3 - 8x_1 + 4x_2 - 8x_3$

Apply one iteration of Newton's method at $x_k = (1, 0, 1)$.

$$f(x) = \frac{1}{2} x^T \begin{bmatrix} 4 & 0 & 4 \\ 0 & -2 & 0 \\ 4 & 0 & 4 \end{bmatrix} x + \begin{bmatrix} -8 & 4 & -8 \end{bmatrix} x$$

$$f'(x) = \nabla f(x)^T = \left(\begin{bmatrix} 4 & 0 & 4 \\ 0 & -2 & 0 \\ 4 & 0 & 4 \end{bmatrix} x + \begin{bmatrix} -8 \\ 4 \\ -8 \end{bmatrix} \right)^T$$

$$f'(x_k) = \begin{bmatrix} 0 & 4 & 0 \end{bmatrix}$$

$$f(x_k) = -8$$

(1) Newton step p_k satisfies

$$f'(x_k)\, p_k = -f(x_k)$$

$$\Longrightarrow \begin{bmatrix} 0 & 4 & 0 \end{bmatrix} p_k = 8$$

$$\Longrightarrow p_k = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} + c \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \text{for all } c \in \mathbb{R}$$

(2) Next estimate (supposing $c = 0$)

$$x_{k+1} = x_k + p_k = \begin{bmatrix} 1 \\ 2 \\ \end{bmatrix}$$

LINEAR APPROXIMATION USED

$$(*) \quad F'(x_k)\, p_k = -F(x_k)$$

where $p_k = x_{k+1} - x_k$.

Consequently $(*)$ can be written as

$$0 = F(x_k) + F'(x_k)(x_{k+1} - x_k),$$

equivalently

$$\boxed{\begin{array}{l} x_{k+1} \text{ is a root of linear approximation about } x_k \\ L_k(x) = F(x_k) + F'(x_k)(x - x_k) \end{array}}$$

EXAMPLE

$F: \mathbb{R}^3 \to \mathbb{R}^3,$

$$F(x) = \begin{bmatrix} x_1^3 + x_3^2 \\ x_3 e^{x_2} \\ \sin x_2 \end{bmatrix}$$

Find the linear approximation for $F(x)$ about $x_k = (1, 0, 1)$ (used by Newton's method.)

Jacobian

$$F'(x) = \begin{bmatrix} 3x_1^2 & 0 & 2x_3 \\ 0 & x_3 e^{x_2} & e^{x_2} \\ 0 & \cos x_2 & 0 \end{bmatrix}$$

$$F'(x_k) = \begin{bmatrix} 3 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Function value

$$F(x_k) = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Linear approximation about x_k

$$L(x) = \underbrace{\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}}_{F(x_k)} + \underbrace{\begin{bmatrix} 3 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}}_{F'(x_k)} \left(x - \underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}_{x_k} \right)$$

$$= \begin{bmatrix} 3(x_1 - 1) + 2(x_3 - 1) + 2 \\ x_2 + (x_3 - 1) + 1 \\ x_2 \end{bmatrix}$$

Next Newton estimate is a root of $L(x)$.

$$L(x_{k+1}) = 0 \implies x_{k+1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

MATRIX NORMS

DEFN (Matrix norm)

A matrix norm $\|\cdot\| : \mathbb{R}^{m\times n} \to \mathbb{R}$ satisfying

(POSITIVITY) (1) $\|A\| > 0$ for all nonzero $A \in \mathbb{R}^{m\times n}$

(HOMOGENEITY) (2) $\|\alpha A\| = |\alpha| \, \|A\|$ for all $\alpha \in \mathbb{R}$ and $A \in \mathbb{R}^{m\times n}$

(TRIANGLE INEQUALITY) (3) $\|A+B\| \leqslant \|A\| + \|B\|$ for all $A, B \in \mathbb{R}^{m\times n}$

Additionally most matrix norms satisfy the submultiplicative property

$$\|AB\| \leqslant \|A\| \, \|B\| \quad \text{for all } A, B$$

Common matrix norms

$\boxed{\text{Frobenius norm}}$

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}$$

e.g.

$$\left\| \begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix} \right\|_2 = \sqrt{(2)^2 + (-3)^2 + (1)^2 + (2)^2}$$

$$= \sqrt{18}$$

PROPERTY

$$\|A\|_F = \sqrt{\operatorname{trace}(A^T A)}$$

where $\operatorname{trace}(B) = \sum_{j=1}^{n} b_{jj}$ for $B \in \mathbb{R}^{n \times n}$.

(Verify as an exercise)

$\boxed{\text{2-norm (or spectral norm)}}$

The maximal amount the mapping $x \longrightarrow Ax$ can stretch a unit vector x

$$\|A\|_2 = \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t} \\ \|x\|_2 = 1}} \|Ax\|_2$$

e.g.

$$\left\| \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \right\| = \max_{\substack{x \in \mathbb{R}^2 \\ \text{s.t.} \\ \|x\|_2 = 1}} \left\| \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_2$$

$$= \max_{\|x\|_2 = 1} \left\| \begin{bmatrix} 2x_1 - x_2 \\ x_1 + 2x_2 \end{bmatrix} \right\|_2$$

$$= \max_{\|x\|_2 = 1} \sqrt{(2x_1 - x_2)^2 + (x_1 + 2x_2)^2}$$

$$= \max_{\|x\|_2 = 1} \sqrt{5x_1^2 + 5x_2^2} = \sqrt{5}$$

1-norm (maximal column sum)

$$\|A\|_1 = \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \\ \|x\|_1 = 1}} \|Ax\|_1$$

THM

$$\|A\|_1 = \max_{1 \leq j \leq n} \|a_j\|_1$$

PROOF

Let $\rho = \|a_k\|_1 = \max_{1 \leq j \leq n} \|a_j\|_1$. Consider an $x \in \mathbb{R}^n$ such that $\|x\|_1 = 1$. We have

$$\begin{aligned}
\|Ax\|_1 &= \|x_1 a_1 + x_2 a_2 + \dots + x_n a_n\|_1 \\
&\leq \|x_1 a_1\|_1 + \|x_2 a_2\|_1 + \dots + \|x_n a_n\|_1 \quad \text{(triang. inequa}\ldots\text{)} \\
&= |x_1|\,\|a_1\|_1 + |x_2|\,\|a_2\|_1 + \dots + |x_n|\,\|a_n\|_1 \quad \text{(homogenei}\ldots\text{)} \\
&\leq \|a_k\|_1 \underbrace{(|x_1| + |x_2| + \dots + |x_n|)}_{\|x\|_1 = 1} \\
&= \|a_k\|_1 = \rho
\end{aligned}$$

$$\Longrightarrow \max_{\|x\|_1 = 1} \|Ax\|_1 \leq \rho$$

- For other direction

$$\max_{\|x\|_1=1} \|Ax\|_1 \geqslant \|A e_k\|_1 = \|a_k\|_1 = \rho$$

where e_k is the kth col of nxn identity matrix

$\square$

e.g.

$$\left\| \begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix} \right\|_2 = \max \left(\left\| \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\|_1, \left\| \begin{bmatrix} -3 \\ 2 \end{bmatrix} \right\|_1 \right)$$

$$= 5$$

$\boxed{\infty\text{-norm (maximal row sum)}}$

$$\|A\|_\infty = \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \\ \|x\|_\infty = 1}} \|Ax\|_\infty$$

<u>THM</u>

Let $A = \begin{bmatrix} \bar{a}_1 \\ \vdots \\ \bar{a}_m \end{bmatrix} \in \mathbb{R}^{m \times n}$ where $\bar{a}_j$ denotes the jth row of A. Then

$$\|A\|_\infty = \max_{1 \leqslant j \leqslant m} \|\bar{a}_j^T\|_1$$

(Proof - exercise)

- e.g.

$$\left\| \begin{bmatrix} 2 & -3 \\ 1 & 2 \end{bmatrix} \right\|_\infty = \max \begin{pmatrix} \| [2 \quad -3]^T \|_1 , \\ \| [1 \quad 2]^T \|_1 \end{pmatrix}$$

$$= 5$$

Matrix p-norm (induced by vector p-norm)

$$\|A\|_p = \max_{\substack{x \in \mathbb{R}^n \\ \text{s.t.} \\ \|x\|_p = 1}} \|Ax\|_p$$