Relaxation method

easy to write, but does it work? for what choice of λ?

$$g(x) = x - \lambda f(x)$$

scalar multiplier... (relaxation coefficient)

Newton's method

$$g(x) = x - \left(Df(x)\right)^{-1} f(x)$$

$(J_f(x))^{-1}$ in the book

a natural compromise between relaxation an Newton's method with n equations/unknowns is

$$g(x) = x - M f(x)$$

Theorem 4.3 *Suppose that $\boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{0}$, and that all the first partial derivatives of $\boldsymbol{f} = (f_1, \dots, f_n)^{\mathrm{T}}$ are defined and continuous in some (open) neighbourhood of $\boldsymbol{\xi}$, and satisfy a condition of strict diagonal dominance at $\boldsymbol{\xi}$; i.e.,*

$$\frac{\partial f_i}{\partial x_i}(\boldsymbol{\xi}) > \sum_{\substack{j=1 \\ j \neq i}}^{n} \left| \frac{\partial f_i}{\partial x_j}(\boldsymbol{\xi}) \right|, \quad i = 1, 2, \dots, n. \tag{4.17}$$

Then, there exist $\varepsilon > 0$ and a positive constant λ such that the relaxation iteration (4.16) converges to $\boldsymbol{\xi}$ for any $\boldsymbol{x}_0$ in the closed ball $\bar{B}_\varepsilon(\boldsymbol{\xi})$ of radius ε, centre $\boldsymbol{\xi}$.

$$Df(\xi) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\xi) & \frac{\partial f_1}{\partial x_2}(\xi) & \cdots & \frac{\partial f_1}{\partial x_n}(\xi) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\xi) & \frac{\partial f_n}{\partial x_2}(\xi) & \cdots & \frac{\partial f_n}{\partial x_n}(\xi) \end{bmatrix}$$

sometimes called diagonal dominance

Want to show g is a contraction

$$g(x) = x - \lambda f(x)$$

scalar

By showing $\|Dg(\xi)\|_\infty < 1$.

Know: $\|I\|_\infty = 1$

$$Dg(\xi) = Dx\Big|_{x=\xi} - \lambda Df(\xi) = I - \underbrace{\lambda Df(\xi)}$$

this needs to reduce the norm of I

$$Dx = \begin{bmatrix} \frac{\partial x_1}{\partial x_1} & \frac{\partial x_1}{\partial x_2} & \cdots & \frac{\partial x_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_n}{\partial x_1} & \frac{\partial x_n}{\partial x_2} & \cdots & \frac{\partial x_n}{\partial x_n} \end{bmatrix} = I$$

Case the diagonal

this is ≤ 1

$$\frac{\partial g_i}{\partial x_i}(\xi) = [Dg(\xi)]_{ii} = 1 - \lambda[Df(\xi)]_{ii} = 1 - \frac{1}{m}[Df(\xi)]_{ii} \geq 0$$

want to choose $\lambda = \frac{1}{[Df(\xi)]_{ii}}$ but λ can't depend on i.

instead set $\lambda = \frac{1}{m}$ where $m = \max\{[Df(\xi)]_{ii} : i = 1, \dots, n\}$

$$\|Dg(\xi)\|_\infty = \max_{i=1}^{n} \sum_{j=1}^{n} \left|\left[Dg(\xi)\right]_{ij}\right| = \max_{i=1}^{n} \sum_{j=1}^{n} \left|\frac{\partial g_i}{\partial x_j}(\xi)\right|$$

$$\sum_{j=1}^{n} \left|\frac{\partial g_i}{\partial x_j}(\xi)\right| = \left|\frac{\partial g_i}{\partial x_i}(\xi)\right| + \sum_{j\neq i} \left|\frac{\partial g_i}{\partial x_j}(\xi)\right|$$

$$= \overbrace{1-\frac{1}{m}\frac{\partial f_i}{\partial x_i}(\xi)}^{\text{know this} \geq 0} + \sum_{j\neq i} \left|\frac{1}{m}\frac{\partial f_i}{\partial x_j}(\xi)\right|$$

note have to remove absolute value before substituting the inequality

$$\boxed{\frac{\partial f_i}{\partial x_i}(\xi)} > \sum_{\substack{j=1\\ j\neq i}}^{n} \left|\frac{\partial f_i}{\partial x_j}(\xi)\right|,$$

$$-\frac{\partial f_i}{\partial x_i}(\xi) < -\sum_{j\neq i} \left|\frac{\partial f_i}{\partial x_j}(\xi)\right|$$

$$< 1 - \frac{1}{m}\cancel{\sum_{j\neq i} \left|\frac{\partial f_i}{\partial x_j}(\xi)\right|} + \cancel{\sum_{j\neq i} \left|\frac{1}{m}\frac{\partial f_i}{\partial x_j}(\xi)\right|}$$

Therefore.

$$\sum_{j=1}^{n} \left|\frac{\partial g_i}{\partial x_j}(\xi)\right| < 1$$

and so

$$\|Dg(\xi)\|_\infty = \max_{i=1}^{n} \sum_{j=1}^{n} \left|\frac{\partial g_i}{\partial x_j}(\xi)\right| < 1$$

Therefore by Theorem 4.2 there is $\varepsilon > 0$ so the iteration $x^{(k+1)} = g(x^{(k)})$ converges for $x^{(0)} \in \overline{B}_\varepsilon(\xi)$.

Newton's method...

$$g(x) = x - \underbrace{\left(Df(x)\right)^{-1}}_{\left(J_f(x)\right)^{-1} \text{ in the book}} f(x)$$

Would like to show that $\| Dg(\xi) \| < 1$.

$$Dg(x) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \end{bmatrix}$$

What is $D\left(\left[Df(x)\right]^{-1} f(x)\right)$? Compute element at a time.

Let $K(x) = \left[Df(x)\right]^{-1}$

definition of matrix multiplication.

$$\left[\left[Df(x)\right]^{-1} f(x)\right]_i = \left[K(x)f(x)\right]_i = \sum_{r=1}^{n} K_{ir}(x) f_r(x)$$

$$\frac{\partial \left[\left[Df(x)\right]^{-1} f(x)\right]_i}{\partial x_j} = \frac{\partial}{\partial x_j} \sum_{r=1}^{n} K_{ir}(x) f_r(x)$$

$$= \sum_{r=1}^{n} \frac{\partial K_{ir}(x)}{\partial x_j} f_r(x) + \sum_{r=1}^{n} K_{ir}(x) \frac{\partial f_r(x)}{\partial x_j}$$

related to second order derivatives...

$$\left.\frac{\partial\left[[Df(x)]^{-1}f(x)\right]_i}{\partial x_j}\right|_{x=\xi} = \sum_{r=1}^{n} \cancel{\frac{\partial K_{ir}(\xi)}{\partial x_j}\underbrace{f_r(\xi)}_{\text{since } f(\xi)=0}} + \sum_{r=1}^{n} K_{ir}(\xi)\frac{\partial f_r(\xi)}{\partial x_j}$$

matrix - matrix mult.

$$= \sum_{r=1}^{n} K_{ir}(\xi)\frac{\partial f_r(\xi)}{\partial x_j} = \left[K(\xi)\, Df(\xi)\right]_{ij}$$

recall $K(x) = [Df(x)]^{-1}$

$$= \left[[Df(\xi)]^{-1} Df(\xi)\right]_{ij} = [I]_{ij}$$

$$g(x) = x - \underbrace{(Df(x))^{-1}}_{(J_f(x))^{-1} \text{ in the book}} f(x)$$

$$Dg(x) = I - D\left((Df(x))^{-1} f(x)\right)$$

$$Dg(\xi) = I - I = 0 = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}$$

$\|Dg(\xi)\| = 0$ for any matrix norm.

In particular...

by Theorem 4.2 there is $\varepsilon > 0$ so the iteration $x^{(k+1)} = g(x^{(k)})$ converges for $x^{(0)} \in \overline{B}_\varepsilon(\xi)$, What is the rate of convergence...

Theorem 4.4 *Suppose that* $\boldsymbol{f}(\boldsymbol{\xi}) = \mathbf{0}$, *that in some (open) neighbourhood* $N(\boldsymbol{\xi})$ *of* $\boldsymbol{\xi}$, *where* $\boldsymbol{f}$ *is defined and continuous, all the second-order partial derivatives of* $\boldsymbol{f}$ *are defined and continuous, and that the Jacobian matrix* $Df(\boldsymbol{\xi})$ *of* $\boldsymbol{f}$ *at the point* $\boldsymbol{\xi}$ *is nonsingular. Then, the sequence* $(\boldsymbol{x}^{(k)})$ *defined by Newton's method (4.18) converges to the solution* $\boldsymbol{\xi}$ *provided that* $\boldsymbol{x}^{(0)}$ *is sufficiently close to* $\boldsymbol{\xi}$; *the convergence of the sequence* $(\boldsymbol{x}^{(k)})$ *to* $\boldsymbol{\xi}$ *is at least quadratic.*

Computer lab on Friday... Check Thursday for more details on the lab.