

Chapter 10 Nonlinear Equations

§ Newton's method

$f_i(\bar{X}) = 0$, $\bar{X} = \{x_1, x_2 \dots, x_n\}^T$, $i = 1 \sim n$, have a solution $\bar{X}^*$.

Expand $f_i(\bar{X})$ about $\bar{X}^*$ into the Taylor's series

$$f_i(\bar{X}) = f_i(\bar{X}^*) + \nabla f_i(\bar{X}^*) \bullet (\bar{X} - \bar{X}^*), \quad i = 1 \sim n$$

$$\bar{F}(\bar{X}) = \bar{F}(\bar{X}^*) + J(\bar{X}^*)(\bar{X} - \bar{X}^*)$$

$$\bar{F}(\bar{X}) = \begin{Bmatrix} f_1 \\ \bullet \\ \bullet \\ f_n \end{Bmatrix}(\bar{X}), \quad J(\bar{X}) = \left[\frac{\partial f_i}{\partial x_j} \right] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \bullet & \bullet & \frac{\partial f_1}{\partial x_n} \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \frac{\partial f_n}{\partial x_1} & \bullet & \bullet & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

Since $\bar{F}(\bar{X}^*) = \{0\}$, $\bar{X}^* = \bar{X} - J^{-1}(\bar{X}^*)\bar{F}(\bar{X})$

$$\bar{X}^{(j+1)} = \bar{X}^{(j)} - J^{-1}(\bar{X}^{(j)})\bar{F}(\bar{X}^{(j)})$$

§ Broyden's method (Quasi-Newton Method)

Modify $\bar{X}^{(j+1)} = \bar{X}^{(j)} - J^{-1}(\bar{X}^{(j)})\bar{F}(\bar{X}^{(j)})$ to $\bar{X}^{(j+1)} = \bar{X}^{(j)} - A_j^{-1}\bar{F}(\bar{X}^{(j)})$

where

$$A_1 = J(\bar{X}^{(0)}) + \frac{[\bar{F}(\bar{X}^{(1)}) - \bar{F}(\bar{X}^{(0)}) - J(\bar{X}^{(0)})(\bar{X}^{(1)} - \bar{X}^{(0)})]}{\|\bar{X}^{(1)} - \bar{X}^{(0)}\|_2^2}(\bar{X}^{(1)} - \bar{X}^{(0)})^T,$$

$$A_i = A_{i-1} + \frac{\bar{Y}^{(i)} - A_{i-1}\bar{S}^{(i)}}{\|\bar{S}^{(i)}\|_2^2} \bar{S}^{(i)T}, \quad A_i^{-1} = A_{i-1}^{-1} + \frac{\bar{S}^{(i)} - A_{i-1}^{-1}\bar{Y}^{(i)}}{\bar{S}^{(i)} A_{i-1}^{-1} \bar{Y}^{(i)}} \bar{S}^{(i)T} A_{i-1}^{-1}$$

in which

$$\bar{Y}^{(i)} = \bar{F}(\bar{X}^{(i)}) - \bar{F}(\bar{X}^{(i-1)}), \quad \bar{S}^{(i)} = \bar{X}^{(i)} - \bar{X}^{(i-1)}$$

§ The Steepest Descent Method

$f_i(\bar{X}) = 0$, $\bar{X} = \{x_1, x_2 \dots, x_n\}^T$, $i = 1 \sim n$, implies

$$g(\bar{X}) = \sum_{i=1}^n f_i^2(\bar{X}) = 0$$

So, the solution $\bar{X}^*$ will minimize $g(\bar{X}) = \sum_{i=1}^n f_i^2(\bar{X}) = 0$

$$g(\bar{X} + d\bar{X}) = g(\bar{X}) + \nabla g(\bar{X}) \circ d\bar{X}$$

Let $d\bar{X} = -\alpha \nabla g(\bar{X}) / |\nabla g(\bar{X})|$, then $g(\bar{X} + d\bar{X}) < g(\bar{X})$

$$\bar{X}^{(k+1)} = \bar{X}^{(k)} - \alpha \nabla g(\bar{X}^{(k)}) / |\nabla g(\bar{X}^{(k)})|$$

How to choose α such that $\min_{\alpha} h(\alpha) = \min_{\alpha} g(\bar{X}^{(k)} - \alpha \nabla g(\bar{X}^{(k)}) / |\nabla g(\bar{X}^{(k)})|)$

It is difficult to find α .

Try $\{\alpha_i, h(\alpha_i)\}_{i=1}^3$, where $\alpha_1 < \alpha_2 < \alpha_3$

Then the Lagrange's interpolation is

$$h(\alpha) = \frac{(\alpha - \alpha_2)(\alpha - \alpha_3)}{(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)} h(\alpha_1) + \frac{(\alpha - \alpha_1)(\alpha - \alpha_3)}{(\alpha_2 - \alpha_1)(\alpha_2 - \alpha_3)} h(\alpha_2) + \frac{(\alpha - \alpha_1)(\alpha - \alpha_2)}{(\alpha_3 - \alpha_1)(\alpha_3 - \alpha_2)} h(\alpha_3)$$

Let $h'(\alpha) = 0$ and find the optimal α^* .

$$\text{So, } \bar{X}^{(k+1)} = \bar{X}^{(k)} - \alpha^* \nabla g(\bar{X}^{(k)}) / |\nabla g(\bar{X}^{(k)})|.$$

§ Continuation methods

The problem $\bar{F}(\bar{x}) = \bar{0}$ has the solution $\bar{x}^*$.

Modify $\bar{F}(\bar{x}) = \bar{0}$ into the form

$$\bar{0} = \bar{G}(\lambda, \bar{x}) = \lambda \bar{F}(\bar{x}) + (1 - \lambda)[\bar{F}(\bar{x}) - \bar{F}(\bar{x}(0))] = \bar{F}(\bar{x}) - (1 - \lambda)\bar{F}(\bar{x}(0))$$

$$, \quad 0 \leq \lambda \leq 1$$

with

$$\text{a. } \bar{0} = \bar{G}(0, \bar{x}) = [\bar{F}(\bar{x}) - \bar{F}(\bar{x}(0))] \quad \text{or} \quad \bar{x} = \bar{x}(0) \quad \text{at} \quad \lambda = 0$$

$$\text{b. } \bar{0} = \bar{G}(1, \bar{x}) = \bar{F}(\bar{x}) \quad \text{or} \quad \bar{x}^* = \bar{x}(1) \quad \text{at} \quad \lambda = 1.$$

$$\begin{aligned}\bar{0} &= \frac{d}{d\lambda} \bar{G}(\lambda, \bar{x}) = \frac{\partial \bar{G}(\lambda, \bar{x})}{\partial \lambda} + \frac{\partial \bar{G}(\lambda, \bar{x})}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial \lambda} \\ \frac{\partial \bar{x}}{\partial \lambda} &= -\left[\frac{\partial \bar{G}(\lambda, \bar{x})}{\partial \bar{x}}\right]^{-1} \frac{\partial \bar{G}(\lambda, \bar{x})}{\partial \lambda} = -\left[\frac{\partial f_i(\lambda, \bar{x})}{\partial x_j}\right]^{-1} \bar{F}(\bar{x}(0)) \\ &= -[J(\bar{x}(\lambda))]^{-1} \bar{F}(\bar{x}(0))\end{aligned}$$

which can be solved by the Runge-Kutta method.

$$1. \quad h = 1/N, \quad \lambda_0 = 0, \quad \bar{x}(\lambda_0) = \bar{x}(0), \quad \bar{w}_0 = \bar{x}(0)$$

$$2. \quad \text{For } j = 0 \sim N$$

$$\lambda_{j+1} = \lambda_j + h$$

$$\bar{k}_1 = h[-J(\bar{w}_{1,j}, \dots, \bar{w}_{n,j})]^{-1} \bar{F}(\bar{x}(\lambda_j)) = h[-J(\bar{w}_j)]^{-1} \bar{F}(\bar{x}(\lambda_j)),$$

$$\bar{k}_2 = h[-J(\bar{w}_j + \frac{1}{2} \bar{k}_1)]^{-1} \bar{F}(\bar{x}(\lambda_j)), \quad \bar{k}_3 = h[-J(\bar{w}_j + \frac{1}{2} \bar{k}_2)]^{-1} \bar{F}(\bar{x}(\lambda_j)),$$

$$\bar{k}_4 = h[-J(\bar{w}_j + \bar{k}_3)]^{-1} \bar{F}(\bar{x}(\lambda_j)),$$

$$\bar{x}(\lambda_{j+1}) = \bar{x}(\lambda_j) + \frac{1}{6}(\bar{k}_1 + 2\bar{k}_2 + 2\bar{k}_3 + \bar{k}_4) = \bar{w}_j + \frac{1}{6}(\bar{k}_1 + 2\bar{k}_2 + 2\bar{k}_3 + \bar{k}_4).$$

HW 10

Using the following methods to solve the problem

$$10x_1 + 2x_2^2 - 3x_3 = 9$$

$$x_1^2 + 3x_2 - x_3^2 = -2$$

$$x_1 x_2 x_3 = 6$$

(a) Newton's method, (b) Broyden's method, (c) the steepest descent method, and (d) the Continuation method $N = 4$.