2006–05–16

Icke linjär minsta kvadrat

$$\min_{\boldsymbol{x}} \|\boldsymbol{f}(\boldsymbol{x})\|_2 \approx \min_{\boldsymbol{x}} \|\boldsymbol{f}(\boldsymbol{x}_k) + \mathcal{J}(\boldsymbol{x} - \boldsymbol{x}_k)\|$$

Sätt $\boldsymbol{d}_k = \boldsymbol{x} - \boldsymbol{x}_k$

$$\Longleftrightarrow \begin{cases} \boldsymbol{x} = \boldsymbol{x}_k + \boldsymbol{d}_k \\ \mathcal{J}(\boldsymbol{x}_k)\, \boldsymbol{d}_k = -\, \boldsymbol{f}(\boldsymbol{x}_k) \end{cases}$$

(Överbestämt linjärt system, löses i minsta-kvadrat-mening.)

Gauss-Newtons metod:

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{d}_k \\ \mathcal{J}(\boldsymbol{x}_k)\, \boldsymbol{d}_k = -\, \boldsymbol{f}(\boldsymbol{x}_k) \end{cases}, \quad k = 0, 1, \ldots$$

Lokalt konvergent.

En populär metod i programvara är Levenberg-Marquardts metod som är *globalt* konvergent (kommer ej på tentan).

$\min_{\boldsymbol{x}} \|\boldsymbol{f}(\boldsymbol{x})\|_2$ approximeras med linjär model med steglängdskontroll.

$$\min_{\boldsymbol{x}} \|\boldsymbol{f}(\boldsymbol{x})\|_2 \approx \min_{\|\boldsymbol{d}\|_2 \leqslant \Delta} \|\boldsymbol{f}(\boldsymbol{x}_k) + \mathcal{J}(\boldsymbol{x}_k)\, \boldsymbol{d}_k\|_2$$

Detta är ekvivalent med (visas inte) att lösa:

$$(\mathcal{J}(\boldsymbol{x}_k)^T \mathcal{J}(\boldsymbol{x}_k) + \mu I)\, \boldsymbol{d} = -\, \mathcal{J}(\boldsymbol{x}_k)^T \boldsymbol{f}(\boldsymbol{x}_k)$$

Man kan dämpa steglängden genom att välja ett stort $\mu > 0$: väljs så att $\|\boldsymbol{d}\|_2 \leqslant \Delta$.

Anmärkning. För $\mu = 0$ är detta normalekvationerna. Men inte så illa konditionerade om $\mu > 0$. Vi gör alltså en reguljärisering.

Levenberg-Marquardts metod är alltså:

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{d}_k \\ (\mathcal{J}(\boldsymbol{x}_k)^T \mathcal{J}(\boldsymbol{x}_k) + \mu I)\, \boldsymbol{d}_k = -\, \mathcal{J}(\boldsymbol{x}_k)^T \boldsymbol{f}(\boldsymbol{x}_k) \end{cases}$$

I praktiken får man pröva sig fram till ett μ så att $\|\boldsymbol{d}\|_2 \leqslant \Delta$, exempelvis genom intervallhalvering.

Tvåpunkts randvärdesproblem, Heath 10.

$$\text{(RVP)} \quad \begin{cases} \alpha\, y'' + \beta\, y' + \gamma\, y = f(t), & 0 \leqslant t \leqslant 1 \\ y(0) = c_1, & y(1) = c_2 \end{cases}$$

$y(t)$ sökt funktion, $f(t), \alpha, \beta, \gamma, c_1, c_2$ är givna.

Metod 1: Differensapproximation av differenskvoterna ger approximation till $y(t)$ i diskreta punkter ⟨fig25⟩. Låt $f_i = f(t_i)$. Centraldifferens ger linjärt ekvationssystem:

$$\begin{cases} \frac{\alpha}{h^2}(y_{i-1} - 2\, y_i + y_{i+1}) + \frac{\beta}{2\, h}(y_{i+1} - y_{i-1}) + \gamma\, y_i = f_i, & i = 2, \ldots, n-1 \\ y_1 = c_1, \quad y_n = c_2 \end{cases}$$

På matrisform, multiplicera med h^2.

$$\begin{pmatrix} \ddots & \ddots & & & & \\ \ddots & \ddots & \ddots & & & \mathbf{0} \\ & \ddots & \ddots & \ddots & & \\ & & \alpha - \frac{\beta h}{2} & -2\alpha + \gamma h^2 & \alpha + \frac{\beta h}{2} & \\ & & & \ddots & \ddots & \ddots \\ & & & & \ddots & \ddots \\ \mathbf{0} & & & & & \end{pmatrix} \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} h^2 f_2 - c_1\left(\alpha - \frac{\beta h}{2}\right) \\ h^2 f_3 \\ \vdots \\ h^2 f_i \\ \vdots \\ h^2 f_{n-2} \\ h^2 f_{n-1} - c_2\left(\alpha + \frac{\beta h}{2}\right) \end{pmatrix}$$

Tridiagonalt, lätt att lösa.

Metod 2, inskjutning. Skriv om (RVP) till (BVP).

$$y_1 = y$$

$$y_2 = y'$$

$$\begin{cases} y_1' = y_2 \\ y_2' = \frac{1}{\alpha}[-\beta y_2 - \gamma y_1 + f(t)] \\ y_1(0) = c_2 \\ y_2(0) = s \end{cases}$$

där s är en okänd parameter. Välj nu s så att lösningen

$$y_1(1, s) = c_2$$

dvs vi löser det givna (RVP) om vi väljer s rätt. Vi ska alltså lösa ekvationen

$$q(s) \equiv y_1(1, s) = c_2$$

Metoder för ekvationslösning får användas.

Exempel. Sekantmetoden: $s_{k+1} = s_k - \frac{[y_1(1, s_k) - c_2](s_k - s_{k-1})}{y_1(1, s_k) - y_2(1, s_{k-1})}, k = 1, 2, \ldots$

Två startvärden (provskott): s_0, s_1.

Exempel. Inskjutning. ⟨fig26⟩.

$$(\text{RVP}) \quad \begin{cases} EI \dfrac{y''}{(1 + (y')^2)^{\frac{3}{2}}} = Py - \frac{w x^2}{2} \\ y(0) = 0 \\ y'(L) = 0 \end{cases}$$

Inskjutning: $y_1 = y, y_2 = y'$:

$$\begin{cases} y_1' = y_2 \\ y_2' = f(x, y, y') \quad \text{från givna problemet} \\ y_1(0) = 0 \\ y_2(0) = s \end{cases}$$

Ekvation: $q(s) = y_2(L, s,) = 0$. Sekantmetoden:

$$s_{k+1} = s_k - \frac{y_2(L, s_k)}{y_2(L, s_k) - y_2(L, s_{k-1})}$$

ANMÄRKNING. I varje iteration av sekantmetoden får vi lösa ODE-problemet med t.ex. `ode45` i MALTAB.

EXEMPEL (tentauppgift) Undersök om algoritmen

$$y = a \cdot x + b$$

är stabil i ett IEEE-system. Vi antar att a, x obh b är lagrade flyttal i datorn.

$$\mathrm{fl}(a \cdot x) = (1 + \varepsilon_1)\,(a \cdot x)$$

där $|\varepsilon_1| \leqslant \mu$.

$$\mathrm{fl}(a \cdot x + b) = (1 + \varepsilon_2)[(1 + \varepsilon_1)(a \cdot x) + b]$$

Är algoritmen stabil? Bakåtfelet ska vara litet. Vi ska hittta $\hat{a}, \hat{x}, \hat{b}$ så att

$$\mathrm{fl}(a \cdot x + b) = \hat{a} \cdot \hat{x} + \hat{b}$$

Vi har

$$\mathrm{fl}(a \cdot x + b) = (1 + \varepsilon_2)\,a \cdot (1 + \varepsilon_1)\,x + (1 + \varepsilon_2)\,b = \hat{a} \cdot \hat{x} + \hat{b}$$

$$\frac{\hat{a} - a}{a} = \varepsilon_2$$

$$\left|\frac{\hat{a} - a}{a}\right| \leqslant \mu$$

På samma sätt för $\hat{x}$ och $\hat{b}$. Algoritmen är stabil.

UPPGIFT: Visa att den asymptotiska felkonstanten $C = \frac{1}{2}$ vid dubbelrot (linjär konvergens) hos Newtons metod.

$$f(x_k) \approx f(x^*) + f'(x^*)\,(x^* - x_k) + \frac{1}{2} f''(x^*)\,(x_k - x^*)^2 = \frac{1}{2} f''(x^*)(x_k - x^*)^2$$

$$f'(x_k) \approx f'(x^*) + f''(x^*)(x^* - x_k)$$

Newtons metod:

$$x_k - \frac{f(x_k)}{f'(x_k)} = x_k - \frac{\frac{1}{2} f''(x^*)(x_k - x^*)^2}{f''(x^*)(x^* - x_k)} + \ldots$$

$$x_{k+1} - x^* = x_k - x^* - \frac{1}{2}(x_k - x^*)$$

$$\frac{x_{k+1} - x^*}{x_k - x^*} \to 1 - \frac{1}{2} = \frac{1}{2} \quad \text{då } k \to \infty$$

Kvasi-Newton

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{s}_k \\ B_k\, s_k = -\,\boldsymbol{f}(\boldsymbol{x}_k) \end{cases}$$

B_k är en approximation av $\mathcal{J}(\boldsymbol{x}_k)$.

Broydens metod:

$$B_{k+1} = B_k + \frac{1}{\|\boldsymbol{s}_k\|_2^2}[\boldsymbol{f}(\boldsymbol{x}_{k+1}) - \boldsymbol{f}(\boldsymbol{x}_k) - B_k\, \boldsymbol{s}_k]\, \boldsymbol{s}_k^T$$

Visa att

1. $B_{k+1}\, \boldsymbol{s}_k = \boldsymbol{f}(\boldsymbol{x}_{k+1}) - \boldsymbol{f}(\boldsymbol{x}_k)$
2. $B_{k+1}\, \boldsymbol{z} = B_k\, \boldsymbol{z}$ då $\boldsymbol{z}^T \boldsymbol{s}_k = 0$

Bevis. 2:

$$B_{k+1}\, \boldsymbol{z} = B_k\, \boldsymbol{z} + \frac{1}{\|\boldsymbol{s}_k\|_2^2}[\boldsymbol{f}(\boldsymbol{x}_{k+1}) - \boldsymbol{f}(\boldsymbol{x}_k) - B_k\, \boldsymbol{s}_k]\, \boldsymbol{s}_k^T \boldsymbol{z} \Longrightarrow B_{k+1}\, \boldsymbol{z} = B_k\, \boldsymbol{z}$$

1:

$$B_{k+1}\, \boldsymbol{s}_k = B_k\, \boldsymbol{s}_k + \frac{1}{\|\boldsymbol{s}_k\|_2^2}[\boldsymbol{f}(\boldsymbol{x}_{k+1}) - \boldsymbol{f}(\boldsymbol{x}_k) - B_k\, \boldsymbol{s}_k]\, \boldsymbol{s}_k^T \boldsymbol{s}_k =$$

$$= B_k\, \boldsymbol{s}_k + \boldsymbol{f}(\boldsymbol{x}_{k+1}) - \boldsymbol{f}(\boldsymbol{x}_k) - B_k \boldsymbol{s}_k = \boldsymbol{f}(\boldsymbol{x}_{k+1}) - \boldsymbol{f}(\boldsymbol{x}_k).$$ □

1. Visa att $\|\boldsymbol{x}\, \boldsymbol{y}^T\|_2 = \|\boldsymbol{x}\|_2 \cdot ||\boldsymbol{y}||_2$ för $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$.

Ledning: Matrisnorm + Cauchys olikhet.

2. Visa att varje skalärprodukt $\langle \cdot, \cdot \rangle \in \mathbb{R}^n$ kan skrivas $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{y}^T A \boldsymbol{x}$ där A är symmetrisk, positivt definit.

Ledning: $\exists$ bas + egenskaper hos skalärprodukt.