NUMPFMF 2025-12-05

[sistemi nelinearnih enačb]

$$f_1(x_1,\dots,x_n)=0 \qquad F=(f_1,\dots,f_n)^T$$
$$\vdots$$
$$f_n(x_1,\dots,x_n)=0 \qquad X=(x_1,\dots,x_n)^T$$

$$\Updownarrow$$

navadna iteracija
izberemo začetni približek $X^{(0)}$

$$F(X)=0 \iff X=G(X)$$
$$X^{(r+1)}=G(X^{(r)})$$

če konvergira; $\alpha=G(\alpha)$; je α rešitev.

njutnova metoda / metoda Newton-Raphson

Jacobijeva matrika

$$X^{(r+1)}=X^{(r)}-J_F^{-1}(X^{(r)})\cdot F(X^{(r)})$$

$$\underbrace{J_F(X^{(r)})}_{A\in\mathbb{R}^{n\times n}}\ \underbrace{\Delta X^{(r)}}_{x}\ \underbrace{=}_{=}\ \underbrace{-F(X^{(r)})}_{b}$$

Problem reševanja nelinearnih enačb pretvorimo na večkratno reševanje linearnih problemov.

Izpeljava njutnove metode

$F(X)=0$. $X^{(r)}$ je približek. Iščemo „popravek" $\Delta X^{(r)}$, da bo $F(X^{(r)}+\Delta X^{(r)})=0$.

$$F(X)=\begin{bmatrix} f_1(X)\\ \vdots \\ f_n(X)\end{bmatrix}$$

razvijmo v Taylorjevo vrsto

$$f_i(X+\Delta X)=f_i(X)+\frac{\partial f_i}{x_1}\Delta x_1+\dots+\frac{\partial f_i}{x_n}\Delta x_n+\cancel{\dots}$$

$$X=(x_1,\dots,x_n)^T$$
$$\Delta X=(\Delta x_1,\dots,\Delta x_n)^T$$

ignoriramo te člene, ker imajo ΔX_i na kvadrat oz. višjo potenco in so zato manjši.

$$f_i:\Omega\subseteq\mathbb{R}^n\to\mathbb{R}$$

Poleg tega pa bi kvadrati predstavljali nelinearnost kar povzroči circular-dependency; ne moremo v reševanju nelin. sist. reševati nelin. sist.

$$F(X+\Delta X)=0\approx F(X)+\begin{bmatrix}\nabla f_1(X)\\ \nabla f_2(X)\\ \vdots\\ \nabla f_n(X)\end{bmatrix}\cdot\Delta X$$

rešujemo

$$0\overset{!}{=}F(X)+J_F(X)\Delta X$$

približek

$$J_F(X)\Delta X=-F(X)$$

iterativno računamo popravke in upamo, da konvergiramo k ničli.

Kako se izognemo računanju odvodov?
često imamo zgolj rešitve.

Namesto $J_F(X^{(r)})$ vzamemo približek za to matriko.

Let B_r približek za $J_F(X^{(r)})$.

a.) reši $B_r \Delta X^{(r)} = -F(X^{(r)})$

b.) $X^{(r+1)} = X^{(r)} + \Delta X^{(r)}$

c.) določi B_{r+1} za naslednji korak.

Broydenova metoda za približke J_F:

- za B_{r+1} vzamemo v spektralni (drugi) normi najbližjo matriko B_r, ki zadošča „SEKANTNEMU POGOJU": $B_{r+1}(X^{(r+1)} - X^{(r)}) = F(X^{(r+1)}) - F(X^{(r)})$.
- za B_0 vzamemo čim boljšo aproksimacijo $J_F(X^{(0)})$, pogosto kar identiteto I.

rešitev (brez dokaza): potegnemo iz rotava

$$B_{r+1} = B_r + \frac{F(X^{(r+1)})\Delta X^{(r)T}}{\Delta X^{(r)T}\Delta X^{(r)}}$$

[Variacijske metode]

$\exists X: \begin{array}{l} f_1(X)=0 \\ \vdots \\ f_n(X)=0 \end{array} \Leftrightarrow G(X)=f_1^2(X)+\dots+f_n^2(X)$ ima globalni minimum enak 0

iščemo minimum fje $G: \Omega \to \mathbb{R}$; $\Omega \subseteq \mathbb{R}^n$

$$G(X) = \sum_{i=1}^{n} f_i^2(X)$$

Potreben pogoj za min.: $\nabla G(X) = 0$

kdaj je to res lokalni minimum?

$$H_G(X) = \begin{bmatrix} G_{x_1x_1}(X) & \cdots & G_{x_1x_n}(X) \\ \vdots & & \vdots \\ G_{x_nx_1}(X) & \cdots & G_{x_nx_n}(X) \end{bmatrix}$$ ← Hessejeva matrika.

če je $H_G(X)$ pozitivno definitna, je pri X lokalni min
če je $H_G(X)$ negativno definitna, je pri X lokalni max
če je $H_G(X)$ nedefinitna, še ne moremo sklepati.

za G vzemimo $G = \sum_{i=1}^{n} f_i^2$.

Splošni pristop:

- naj bo $X^{(r)}$ „tekoči" približek
  - izberemo smer $V_r \in \mathbb{R}^n$ in $\lambda_r \in \mathbb{R}$ tako, da bo vrednost G v novem približku $X^{(r+1)} = X^{(r)} + \lambda_r V_r$ manjša.

želimo $G(X^{(r+1)}) < G(X^{(r)})$. vprašanje je, kako izbrati V_r in λ_r. Sledi nekaj osnovnih idej.

Možnosti: (smer)

a.) Splošna metoda spusta: Izberemo poljubno smer V_r, ki ni pravokotna na $\nabla G(X^{(r)})$.

b.) Metoda najhitrejšega spusta: $V_r = -\nabla G(X^{(r)})$

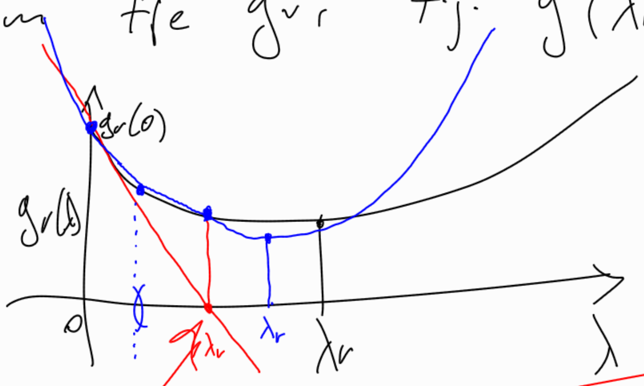
c.) Metoda koordinatnega spusta: za smeri ciklično izbiramo t.j. koordinatne smeri $e_1, e_2, \dots, e_n$.

Kako določimo λ_r?

gledamo f-jo $g_r(\lambda) = G(X^{(r)} + \lambda V_r)$ - f-ja ene spremenljivke in določimo λ_r tako, da je $g_r(\lambda_r) < g_r(0)$

Variante: (izbira λ_r)

a.) Metoda največjega spusta: λ_r določimo tako, da poiščemo minimum f-je g_r, t.j. $g'(\lambda_r) = 0$.



D.N.: kako izračunamo $g'(\lambda)$.

b.) Metoda tangentnega spusta: za λ_r vzamemo presek tangente na g_r z „x-osjo"

c.) Metoda paraboličnega spusta: s tangentno metodo določimo $\alpha = \lambda_r$, nato pa skozi točke

$(0, g_r(0)), (\alpha/2, g_r(\frac{\alpha}{2})), (\alpha, g_r(\alpha))$ potegnemo parabolo. Nato pa za λ_r vzamemo tisto vrednost pri kateri parabola doseže minimum

[Linearni problem najmanjših kvadratov]

Primer: Recimo, da merimo hitrost objekta ob različnih časih t_i; $i=1,2,\dots,N$. Recimo, da vemo, da je hitrost oblike $v(t)=\alpha t+\beta$. Kako določimo α, β? Meritev je zelo veliko in $N \gg 1$. Imamo torej ogromno enačb za dve neznanki, torej problem

$$
\begin{aligned}
\alpha t_1 + \beta &= v_1 \\
\alpha t_2 + \beta &= v_2 \\
&\vdots \\
\alpha t_N + \beta &= v_N
\end{aligned}
$$

Problem skoraj gotovo nima točne rešitve.

rešujemo $Ax=b$, $A \in \mathbb{R}^{m\times n}$, $m \gg n$, $x\in\mathbb{R}^n$, $b\in\mathbb{R}^m$

$$\boxed{A}\,\boxed{x} = \boxed{b}$$

→ temu pravimo rešitev po metodi najmanjših kvadratov

$$x^* = \arg\min \|Ax-b\|_2$$

Vprašanja: ali x^* obstaja, ali je enoličen, kako ga dobimo, ...?

Če A ni polnega ranga ($\operatorname{rang} A < n$), potem rešitev, če obstaja, ni enolična.

$$\operatorname{rang} A < n \Rightarrow \dim \operatorname{Ker} A > 0 \Rightarrow$$

$$\exists\, z \in \operatorname{Ker} A,\ z \neq 0 \ni: Az = 0$$

Recimo, da je x^* opt. rešitev po metodi najmanjših kvadratov. Tedaj je rešitev tudi: $\hat{x} = x^* + \lambda z$; $\lambda \in \mathbb{R}$

$$A\hat{x} - b = A(x^* + \lambda z) - b = Ax^* - b$$

Od tod zato zahtevamo, da je A polnega ranga.

$\Updownarrow$

$\operatorname{rang} A = n$