# 2006–05–02

**System av icke-linjära ekvationer,** Heath 5.6

Problem:

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}, \quad \boldsymbol{f}: \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

DEFINITION: En lösning, en rot $\boldsymbol{x}^*$, är **singulär** ("och det är inte bra") om Jakobianen $J(\boldsymbol{x}^*) = \left( \frac{\partial f_i}{\partial x_j}(\boldsymbol{x}^*) \right)_{i,j}$. Annars är roten **reguljär**.

EXEMPEL:

$$\begin{cases} x_1^3 - x_2 - 1 = 0 \\ x_1^2 - x_2 - 1 = 0 \end{cases}$$

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} x_1^3 - x_2 - 1 \\ x_1^2 - x_2 - 1 \end{pmatrix}, \quad J(\boldsymbol{x}) = \begin{pmatrix} 3\,x_1^2 & -1 \\ 2\,x_1 & -1 \end{pmatrix}$$

$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ har lösningarna

$$\boldsymbol{x}^* = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad \text{och} \quad \boldsymbol{x}^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$J(\boldsymbol{x})$ är singulär i $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$:

$$J\begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 0 & -1 \end{pmatrix}$$

$J(\boldsymbol{x})$ är reguljär i $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$:

$$J\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 & -1 \\ 2 & -1 \end{pmatrix}$$

**Newtons metod**

Analytiskt: linjäriseringen med Taylorutveckling. (Geometriskt: funktionsytor approximeras med tangentplan.)

Taylorutveckling kring approximationen $\boldsymbol{x}_k$:

$$\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{f}(\boldsymbol{x}_k) + J(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{x}_k)$$

En linjär modell av problemet $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ är alltså

$$\boldsymbol{f}(\boldsymbol{x}_k) + J(\boldsymbol{x}) \cdot (\boldsymbol{x} - \boldsymbol{x}_k) = \boldsymbol{0}$$

Detta är ett linjärt ekvationssystem. Lösningen blir vår nya approximation $\boldsymbol{x}_{k+1}$:

$$J(\boldsymbol{x}_k)(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) = -\,\boldsymbol{f}(\boldsymbol{x}_k)$$

Med $\boldsymbol{s}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$ får vi metoden:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{s}_k, \quad J(\boldsymbol{x}_k)\,\boldsymbol{s}_k = -\,\boldsymbol{f}(\boldsymbol{x}_k)$$

Detta är Newtons metod.

Detta är ett exempel på en så kallad **sökmetod**, $\boldsymbol{s}_k$ är **sökriktning**. I Newtons metod bestäms alltså sökriktningen genom ekvationssystemet $J(\boldsymbol{x}_k)\ \boldsymbol{s}_k = -\ \boldsymbol{f}(\boldsymbol{x}_k)$. I varje iteration löser vi ett linjärt ekvationssystem. Newtons metod är **lokalt konvergent**. $\boldsymbol{x}_0$ måste vara tillräckligt nära $\boldsymbol{x}^*$. Konvergensen är kvadratisk vid reguljär rot.

**Varianter av Newtons metod**

1. Dämpad Newton. Syfte: att vidga konvergensområdet, dvs starta längre ifrån $\boldsymbol{x}^*$.

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \alpha_k\, \boldsymbol{s}_k$$

$a_k$ kallas för **steglängd** i sökmetoden.

I vanlig Newton är $\alpha_k = 1$, i dämpad är $\alpha_k < 1$ med villkor $\|\boldsymbol{f}_{k+1}\| < \|\boldsymbol{f}_k\|$.

Närmare lösningen $\boldsymbol{x}^*$ kan vi gå över till $\alpha_k = 1$, dvs vanlig Newton. $\langle$fig17$\rangle$.

2. Modifierad Newton. Syfte: Undvika derivataberäkning och faktorisering av Jakobian.

$J(\boldsymbol{x}_0)$ eller en approximation av $J(\boldsymbol{x}_0)$ används hela tiden, dvs

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{s}_k \\ J(\boldsymbol{x}_0)\, \boldsymbol{s}_k = -\ \boldsymbol{f}(\boldsymbol{x}_k) \end{cases}$$

Detta innebär LU-faktorisering **en** gång av $J(\boldsymbol{x}_0)$, sedan endast lösning av två triangulära system med framåt och bakåtsubstitution i varje iteration.

Första iteration: $\mathcal{O}(n^3)$, men följande iterationer är $\mathcal{O}(n^2)$.

3. Kvasi-Newton-metoder. Syfte: Undvika beräkning av Jakobianen.

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

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

Speciellt: $B_k = J(\boldsymbol{x}_k)$: vanlig Newton. $B_k = J(\boldsymbol{x}_0)$: modifierad Newton.

Idé: $B_{k+1}$ uppdateras från $B_k$ så att $B_{k+1}$ approximerar $J(\boldsymbol{x}_{k+1})$ enligt:

i. $B_{k+1}\ \boldsymbol{s}_k \approx J(\boldsymbol{x}_{k+1})\ \boldsymbol{s}_k$, dvs bra approximation i sökriktningen $\boldsymbol{s}_k$. Detta fixar vi med Taylor:

$$J(\boldsymbol{x}_{k+1})\, \boldsymbol{s}_k = J(\boldsymbol{x}_{k+1})(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) \approx \boldsymbol{f}(\boldsymbol{x}_{k+1}) - \boldsymbol{f}(\boldsymbol{x}_k)$$

som ger oss det så kallade **sekantvillkoret**:

$$B_{k+1}\, \boldsymbol{s}_k = \boldsymbol{f}(\boldsymbol{x}_{k+1}) - \boldsymbol{f}(\boldsymbol{x}_k) \tag{1}$$

ii. $B_{k+1}$ oförändrad i riktningar vinkelräta mot $\boldsymbol{s}_k$.

$$B_{k+1}\, \boldsymbol{z} = B_k\, \boldsymbol{z}, \quad \forall \boldsymbol{z}\colon \boldsymbol{z}^T \boldsymbol{s}_k = 0$$

Övning: Visa att följande uppdatering (Broyden) satisfierar (1) och (2):

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

**Fixpunktsiteration**

Problemet $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ skrivs om på formen $\boldsymbol{x} = \boldsymbol{g}(\boldsymbol{x})$. Sedan iteration enligt $\boldsymbol{x}_{k+1} = \boldsymbol{g}(\boldsymbol{x}_k)$, $k = 0$, $1, \ldots$.

ANMÄRKNING: En punkt $\boldsymbol{x}^*$ sådan att $\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{x}^*$ kallas **fixpunkt** till funktionen $\boldsymbol{g}$. Om fixpunktsiterationen konvergerar så mot en fixpunkt till $\boldsymbol{g}$. Fixpunktsiterationen konvergerar om

$$\|G(\boldsymbol{x})\| \leqslant \mu < 1$$

för $\boldsymbol{x}$ i en omgivning av $\boldsymbol{x}^*$ och $\boldsymbol{x}_0$ ligger i den omgivningen. Här är $G(\boldsymbol{x})$ Jakobianen till $\boldsymbol{g}(\boldsymbol{x})$. Feluppskattning: $\|\boldsymbol{x}_{k+1} - \boldsymbol{x}^*\| \leqslant \frac{\mu}{1-\mu} \|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|$. Beviset ingår inte. För endimensionella fallet, se boken.

### Metodoberoende feluppskattning

Låt $\hat{\boldsymbol{x}}$ vara en approximation till $\boldsymbol{x}^*$ med fel $\delta\boldsymbol{x} = \hat{\boldsymbol{x}} - \boldsymbol{x}^*$. Då med Taylor:

$$\boldsymbol{f}(\hat{\boldsymbol{x}}) = \boldsymbol{f}(\boldsymbol{x}^* + \delta\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}^*) + J(\boldsymbol{x}^*)\,\delta\boldsymbol{x} + \mathcal{O}(\|\delta\boldsymbol{x}\|^2) =$$

$$= J(\boldsymbol{x}^*)\,\delta\boldsymbol{x} + \mathcal{O}(\|\delta\boldsymbol{x}\|^2)$$

$$\boldsymbol{f}(\hat{\boldsymbol{x}}) \approx J(\boldsymbol{x}^*)\,\delta\boldsymbol{x}$$

eller $\delta\boldsymbol{x} = J(\boldsymbol{x}^*)^{-1}\,\boldsymbol{f}(\hat{\boldsymbol{x}})$ med feluppskattning

$$\|\delta\boldsymbol{x}\| \lesssim \|J(\boldsymbol{x}^*)^{-1}\| \cdot \|\boldsymbol{f}(\hat{\boldsymbol{x}})\| \approx \|J(\hat{\boldsymbol{x}})^{-1}\| \cdot \|\boldsymbol{f}(\hat{\boldsymbol{x}})\|$$

Besvärlig att använda i praktiken. Det är dyrt att beräkna inversen.

EXEMPEL:

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} x_1^2 + x_2^2 - 1 \\ \mathrm{e}^{x_1 x_2} + x_1 + x_2 - 1 \end{pmatrix}$$

Lös $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$.

$$J(\boldsymbol{x}) = \begin{pmatrix} 2\,x_1 & 2\,x_2 \\ x_2\,\mathrm{e}^{x_1 x_2} + 1 & x_1\,\mathrm{e}^{x_1 x_2} + 1 \end{pmatrix}$$

Newtons metod: $\boldsymbol{x}_0 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. MATLAB: `x = x - J(x) \ f(x)`.

$$\boldsymbol{x}_1 = \begin{pmatrix} 0.9038 \\ -0.5104 \end{pmatrix}$$

$$\boldsymbol{x}_2 = \begin{pmatrix} 0.8625 \\ -0.5078 \end{pmatrix}$$

$$\boldsymbol{x}_3 = \begin{pmatrix} 0.8617 \\ -0.5075 \end{pmatrix}$$

Med $\hat{\boldsymbol{x}} = \boldsymbol{x}_3$ får vi feluppskattningen $\|\delta\boldsymbol{x}\| \lesssim \|J(\hat{\boldsymbol{x}})^{-1}\| \cdot \|\boldsymbol{f}(\hat{\boldsymbol{x}})\|$. MATLAB:

$$\|J(\hat{\boldsymbol{x}})^{-1}\|_2 = 0.6047$$

$$\left[\text{norm för matris } A: \quad \|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}\right]$$

$$\|\boldsymbol{f}(\hat{\boldsymbol{x}})\|_2 = 8.858 \cdot 10^{-5}$$

$$\Longrightarrow \|\delta\boldsymbol{x}\|_2 \lesssim 5.36 \cdot 10^{-5}$$

**Fixpunktsiteration**

Lös ut $x_1$ ur första ekvationen: $x_1 = (\pm)\sqrt{1-x_2^2}$ och $x_2$ ur andra ekvationen

$$x_2 + \alpha x_2 - \alpha\, x_2 = 1 - x_1 - \mathrm{e}^{x_1 x_2}$$

$$x_2 = \frac{\alpha - 1}{\alpha} x_2 + \frac{1 - x_1 - \mathrm{e}^{x_1 x_2}}{\alpha}$$

Trixandet med $\alpha$ behövs för att åstadkomma konvergens.

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{pmatrix} \sqrt{1-x_2^2} \\ \frac{x_2(\alpha-1) + 1 - x_1 - \mathrm{e}^{x_1 x_2}}{\alpha} \end{pmatrix}$$

Med

$$G(\boldsymbol{x}) = \begin{pmatrix} 0 & -x_2(1-x_2^2)^{-\frac{1}{2}} \\ \frac{-1 - x_2\,\mathrm{e}^{x_1 x_2}}{\alpha} & \frac{\alpha - 1 - x_1\,\mathrm{e}^{x_1 x_2}}{\alpha} \end{pmatrix}$$

$\alpha = 1 \Longrightarrow \|G(\boldsymbol{x}^*)\|_\infty = 1.2287$ ej konvergent.

$\alpha = 2 \Longrightarrow \|G(\boldsymbol{x}^*)\|_\infty = 0.5890$ konvergent.

$$\boldsymbol{x}_0 = \begin{pmatrix} 0.7071 \\ -0.7071 \end{pmatrix}$$

$$\boldsymbol{x}_1 = \begin{pmatrix} 0.7071 \\ -0.5104 \end{pmatrix}$$

$$\vdots$$

$$\boldsymbol{x}_{12} = \begin{pmatrix} 0.8617 \\ -0.5075 \end{pmatrix}$$

**Feluppskattning för fixpunktsiteration**

$$\mu = 0.5890$$

$$\|\boldsymbol{x}_{12} - \boldsymbol{x}^*\| \leqslant \frac{\mu}{1-\mu}\|\boldsymbol{x}_{12} - \boldsymbol{x}_{11}\| \Longrightarrow \|\boldsymbol{x}_{12} - \boldsymbol{x}^*\| \leqslant \frac{0.589}{1-0.589} \cdot 4.10 \cdot 10^{-5} \leqslant 5.9 \cdot 10^{-5}$$