NTNU

Norges teknisk-naturvitenskapelige universitet

Newtons flervariabel metode

Jonas J. Harang

27. oktober 2022

Newtons flervariabel metode

Vi vil løse et ligningssystem:

$$f_1(u,v,w) = 0$$
$$f_2(u,v,w) = 0$$
$$f_3(u,v,w) = 0$$

Det vil si at vi vil finne roten av vektorfunksjonen:

$$F(\mathbf{x}) = \begin{bmatrix} f_1(u,v,w) \\ f_2(u,v,w) \\ f_3(u,v,w) \end{bmatrix}$$

Vi kan gå fram på samme måte som Newtons metode med 1 variabel:

1. Start med en verdi $\mathbf{x_0} = \begin{bmatrix} u_0 & v_0 & w_0 \end{bmatrix}^T$
2. Finn roten $\mathbf{r} = \mathbf{x_1}, \quad F_L(\mathbf{r}) = 0$ av det lineariserte systemet F_L rundt $\mathbf{x_0}$
3. Gjenta for $\mathbf{x_1}$ til ønskelig nøyaktighet er oppnådd

Newtons flervariabel metode

Med 1 variabel blir lineariseringen rundt x_0:

$$f_L(x) = f(x_0) + f'(x) \cdot (x - x_0)$$

Når $F(\mathbf{x}) = \begin{bmatrix} f_1(u,v,w) \\ f_2(u,v,w) \\ f_3(u,v,w) \end{bmatrix}$

må vi ta i bruk *Jacobimatrisen*:

$$DF(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} & \frac{\partial f_1}{\partial w} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} & \frac{\partial f_2}{\partial w} \\ \frac{\partial f_3}{\partial u} & \frac{\partial f_3}{\partial v} & \frac{\partial f_3}{\partial w} \end{bmatrix}$$

- $DF(\mathbf{x})$ er analog med den deriverte $f'(x)$
- $f'(x_0) \cdot x$ gir $f(x)$ dersom $f(x)$ er en rett linje og vi flytter origo til x_0
- $DF(\mathbf{x_0}) \cdot \mathbf{x}$ gir $F(\mathbf{x})$ dersom $F: V \to U$ er en lineær transformasjon og vi flytter origo til $\mathbf{x_0}$

Taylorutvidelse rundt x_0

Taylorutvidelsen av $f(\mathbf{x}) = \begin{bmatrix} f_1(x,y,z) \\ f_2(x,y,z) \\ f_3(x,y,z) \end{bmatrix}$ rundt $\mathbf{x}_0$:

$$
\begin{aligned}
F(\mathbf{x}) = f(\mathbf{x}_0) + \frac{\partial f(\mathbf{x}_0)}{\partial x}(x - x_0) + \frac{\partial f(\mathbf{x}_0)}{\partial y}(y - y_0) \\
+ \frac{\partial f(\mathbf{x}_0)}{\partial z}(z - z_0) + \mathcal{O}((x - x_0)^2, (y - y_0)^2, (z - z_0)^2)
\end{aligned}
$$

Vi finner en *linjær* approksimasjon $F_L(\mathbf{x})$ ved å droppe feilleddet:

$$
\begin{aligned}
F_L(\mathbf{x}) &= f(\mathbf{x}_0) + \frac{\partial f(\mathbf{x}_0)}{\partial x}(x - x_0) + \frac{\partial f(\mathbf{x}_0)}{\partial y}(y - y_0) + \frac{\partial f(\mathbf{x}_0)}{\partial z}(z - z_0) \\
&= f(\mathbf{x}_0) + \begin{bmatrix} \frac{\partial f_1}{\partial x} \\ \frac{\partial f_2}{\partial x} \\ \frac{\partial f_3}{\partial x} \end{bmatrix} (x - x_0) + \begin{bmatrix} \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial y} \\ \frac{\partial f_3}{\partial y} \end{bmatrix} (y - y_0) + \begin{bmatrix} \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial z} \end{bmatrix} (z - z_0) \\
&= f(\mathbf{x}_0) + \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix}
\end{aligned}
$$

Newtons flervariabel metode

Med flere variabler blir linealiseringen rundt $\mathbf{x_0}$:

$$
\begin{aligned}
F(\mathbf{x}) \approx F_L(\mathbf{x}) &= F(\mathbf{x_0}) + DF(\mathbf{x_0}) \cdot (\mathbf{x} - \mathbf{x_0}) \\
F_L(\mathbf{x_1}) = 0 &= F(\mathbf{x_0}) + DF(\mathbf{x_0}) \cdot (\mathbf{x_1} - \mathbf{x_0}) \\
DF(\mathbf{x_0}) \cdot \mathbf{x_1} &= DF(\mathbf{x_0}) \cdot \mathbf{x_0} - F(\mathbf{x_0}) \\
\mathbf{x_1} &= \mathbf{x_0} - (DF(\mathbf{x_0}))^{-1} \cdot F(\mathbf{x_0})
\end{aligned}
$$

- Start med en gjetning $\mathbf{x_0}$ for roten av $F(\mathbf{x})$
- Finn roten av lineariseringen av F rund $\mathbf{x_0} \to \mathbf{x_1}$
- Gjenta til $F(\mathbf{x_n}) \approx 0$

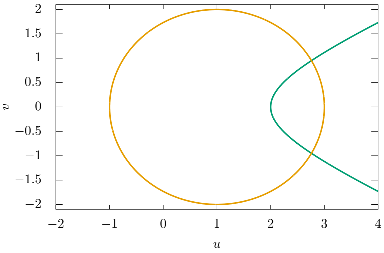
Eksempel 1

Ta to steg med Newtons metode med startpunkt (1,1) for å løse systemet:

$$u^2 - 4v^2 = 4$$
$$(u-1)^2 + v^2 = 4$$

$$F(\mathbf{x}) = \begin{bmatrix} f_1(u,v) \\ f_2(u,v) \end{bmatrix} = \begin{bmatrix} u^2 - 4v^2 - 4 \\ (u-1)^2 + v^2 - 4 \end{bmatrix}$$

$$DF(\mathbf{x}) = \begin{bmatrix} 2u & -8v \\ 2u-2 & 2v \end{bmatrix}$$



$$DF(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial u} & \frac{\partial f_1}{\partial v} \\ \frac{\partial f_2}{\partial u} & \frac{\partial f_2}{\partial v} \end{bmatrix}$$

Eksempel 1, forts

$$\begin{aligned} u^2 - 4v^2 &= 4 \\ (u-1)^2 + v^2 &= 4 \end{aligned} \qquad F(\mathbf{x}) = \begin{bmatrix} u^2 - 4v^2 - 4 \\ (u-1)^2 + v^2 - 4 \end{bmatrix} \qquad DF(\mathbf{x}) = \begin{bmatrix} 2u & -8v \\ 2u-2 & 2v \end{bmatrix}$$

$$\begin{aligned} \mathbf{x_0} &= \begin{bmatrix} 1 & 1 \end{bmatrix}^T \\ F(\mathbf{x_0}) &= \begin{bmatrix} 1-4-4 \\ 0^2+1-4 \end{bmatrix} = \begin{bmatrix} -7 \\ -3 \end{bmatrix} \\ DF(\mathbf{x_0}) &= \begin{bmatrix} 2 & -8 \\ 0 & 2 \end{bmatrix} \\ (DF(\mathbf{x_0}))^{-1} &= \frac{1}{2 \cdot 2 - (-8) \cdot 0} \begin{bmatrix} 2 & 8 \\ 0 & 2 \end{bmatrix} \\ (DF(\mathbf{x_0}))^{-1} &= \begin{bmatrix} 0.5 & 2 \\ 0 & 0.5 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \mathbf{x_1} &= \mathbf{x_0} - (DF(\mathbf{x_0}))^{-1} \cdot F(\mathbf{x_0}) \\ \mathbf{x_1} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0.5 & 2 \\ 0 & 0.5 \end{bmatrix} \cdot \begin{bmatrix} -7 \\ -3 \end{bmatrix} \\ \mathbf{x_1} &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -9.5 \\ -1.5 \end{bmatrix} = \begin{bmatrix} 10.5 \\ 2.5 \end{bmatrix} \end{aligned}$$

Eksempel 1, forts

$$u^2 - 4v^2 = 4 \qquad (u-1)^2 + v^2 = 4$$

$$F(\mathbf{x}) = \begin{bmatrix} u^2 - 4v^2 - 4 \\ (u-1)^2 + v^2 - 4 \end{bmatrix} \qquad DF(\mathbf{x}) = \begin{bmatrix} 2u & -8v \\ 2u-2 & 2v \end{bmatrix}$$

$$\mathbf{x_1} = \begin{bmatrix} 10.5 & 2.5 \end{bmatrix}^T$$

$$F(\mathbf{x_1}) = \begin{bmatrix} 10.5^2 - 4 \cdot 2.5^2 - 4 \\ 9.5^2 + 2.5^2 - 4 \end{bmatrix} = \begin{bmatrix} 81.25 \\ 92.5 \end{bmatrix}$$

$$DF(\mathbf{x_1}) = \begin{bmatrix} 2 \cdot 10.5 & -8 \cdot 2.5 \\ 2 \cdot 10.5 - 2 & 2 \cdot 2.5 \end{bmatrix} = \begin{bmatrix} 21 & -20 \\ 19 & 5 \end{bmatrix}$$

$$(DF(\mathbf{x_1}))^{-1} = \frac{1}{21 \cdot 5 + 20 \cdot 19} \begin{bmatrix} 5 & 20 \\ -19 & 21 \end{bmatrix}$$

$$(DF(\mathbf{x_1}))^{-1} = \begin{bmatrix} 0.0103.. & 0.0412.. \\ -0.0392.. & 0.0433.. \end{bmatrix}$$

$$\mathbf{x_2} = \mathbf{x_1} - (DF(\mathbf{x_1}))^{-1} \cdot F(\mathbf{x_1})$$

$$\mathbf{x_2} = \begin{bmatrix} 10.5 \\ 2.5 \end{bmatrix} - \begin{bmatrix} 4.652.. \\ 0.8222... \end{bmatrix} = \begin{bmatrix} 5.848.. \\ 1.678... \end{bmatrix}$$

$$F(\mathbf{x_2}) = \begin{bmatrix} 18.93.. \\ 22.32.. \end{bmatrix}$$

$$\vdots$$

Mindre kostbar fremgangsmåte

- Å regne ut $(DF(\mathbf{x}))^{-1}$ er kostbart
- Vi kan gjøre følgende forandring

$$
\begin{aligned}
F_L(\mathbf{x_{i+1}}) &= 0 = F(\mathbf{x_i}) + DF(\mathbf{x_i}) \cdot (\mathbf{x_{i+1}} - \mathbf{x_i}) \\
\mathbf{x_{i+1}} - \mathbf{x_i} &= \mathbf{s} \\
-F(\mathbf{x_i}) &= DF(\mathbf{x_i}) \cdot \mathbf{s}
\end{aligned}
$$

Ligningen over kan vi løse for $\mathbf{s}$ ved feks gauss-eliminasjon og tilbakesubstitusjon, og deretter sette $\mathbf{x_{i+1}} = \mathbf{s} + \mathbf{x_i}$

NTNU

Newtons flervariabel metode

```python
import numpy as np
def Newtons_multivariate_method(f, df, x0, tol):
    x = x0
    while(np.abs(np.linalg.norm(f(x))) > tol):
        fx = f(x)
        dfx = df(x)
        s = np.linalg.solve(dfx,-fx)
        x = x+s
    return x

f = lambda x: np.array([x[0]**2-4.0*x[1]**2-4.0,
                        (x[0]-1.0)**2+x[1]**2-4.0]).reshape((2,1))
df = lambda x: np.array([2.0*x[0], -8.0*x[1],
                        2.0*x[0]-2.0, 2.0*x[1]]).reshape((2,2))
x0 = np.array([1,1]).reshape((2,1))
x = Newtons_multivariate_method(f,df,x0,0.001)
print(f"x:\n {x}")
print(f"feil:\n {np.abs(np.linalg.norm(f(x)))}")
```

Eksempeloppgaver fra forrige time

- Gitt en matrise A - Vis at den er SPD ved å:
  - vise at $v^T Av$ alltid er > 0
  - vise til at alle egenverdier er positive
  - Bruke at den er en prinsipal undermatrise av en annen SPD matrise
- Gitt en SPD matrise - finn Cholesky-faktoriseringen
- Gitt et sett med vektorer $V = \{v_1, v_2, ..\}$ og en SPD matrise A - Er vektorene ortogonale? A-ortogonale?
- Gitt en 2×2 SPD matrise A og et sett med konjugerte basisvektorer $P = \{p_1, p_2\}$ bruk $(,)_A$ til å løse $Ax = b$ (Slide 17)
- Gitt en matrise A, vurder hvor godt konjugerte gradienters vil funke (A må være SPD, høyt kondisjonstall, parallelle kolonnevektorer, er ikke bra)
- Hva kan man gjøre dersom A er dårlig kondisjonert (preconditioning)
- Forklar i grove trekk hvordan konjugerte gradienters metode fungerer - Stikkord: SPD, indreprodukt, konjugert basis, krylov-underrom, Gram-Schmidt ...