Newtons metode (seksjon 5.6)

Forrige gang: Fikspunkter (likevektspunkter): $\vec{F}(\vec{x}) = \vec{x}$

I dag: Nullpunkter: $\vec{F}(\vec{x}) = \vec{0}$

Eksempel. Løsning av ligningssystemer:

$$\left.\begin{aligned} x^3 - 3xy^2 &= -1 \\ 3x^2y - y^3 &= 0 \end{aligned}\right\} \text{ to ligninger, to ukjente.}$$

$$\left.\begin{aligned} x^3 - 3xy^2 + 1 &= 0 \\ 3x^2y - y^3 &= 0 \end{aligned}\right\} \text{ omforming med bare 0'er på høyre side}$$

Innfører vi funksjonen $\vec{F}: \mathbb{R}^2 \to \mathbb{R}^2$ ved

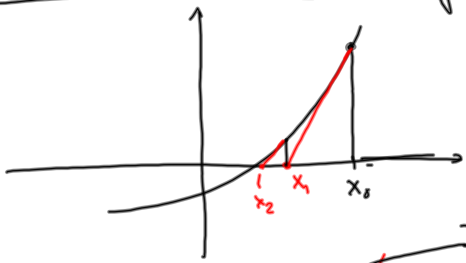
$$\vec{F}(x,y) = \begin{pmatrix} x^3 - 3xy^2 + 1 \\ 3x^2y - y^3 \end{pmatrix},$$ så er det å løse ligningssystemet

det samme som å finne et nullpunkt for $\vec{F}$.

Tilsvarende kan det å løse et ligningssystem med n ligninger og n ukjente omformes til et spørsmål om å finne nullpunkter for en funksjon $\vec{F}: \mathbb{R}^n \to \mathbb{R}^n$.

Newtons metode i én variabel: $f: \mathbb{R} \to \mathbb{R}$, søker $f(x) = 0$

Tipper en tilnærmet løsning x_0.
Ønsker å forbedre tipset.



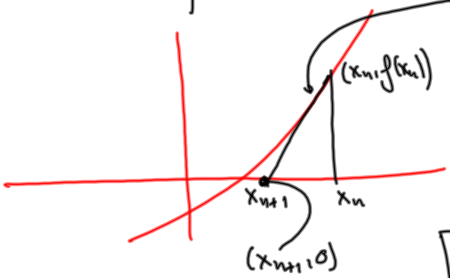
Tangentligningen:

$$y - f(x_n) = f'(x_n)(x - x_n)$$

$$0 - f(x_n) = f'(x_n)(x_{n+1} - x_n)$$

$$-\frac{f(x_n)}{f'(x_n)} = x_{n+1} - x_n$$

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$



Håp: $x_n \to$ nullpunkt.

Newtons metode i flere variable: $\vec{F}: \mathbb{R}^m \to \mathbb{R}^m$, vil finne et nullpunkt $\vec{F}(\vec{x}) = \vec{0}$

Samme framgangsmåte som før: Tipper først en løsning $\vec{x}_0$. Ønsker å finne stadig bedre tips $\vec{x}_0, \vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n, \ldots$. Anta at vi har kommet til $\vec{x}_n$ og ønsker å forbedre dette tipset ytterligere. I prinsippet ønsker vi da å finne en løsning av $\vec{F}(\vec{x}) = 0$ i nærheten av $\vec{x}_n$. En enklere ligning å løse er $T_{\vec{x}_n}\vec{F}(\vec{x}) = \vec{0}$ der $T_{\vec{x}_n}\vec{F}$ er lineariseringen til $\vec{F}$ i $\vec{x}_n$.

Linearisering: $\underbrace{T_{\vec{a}}\vec{F}(\vec{x})}_{\mathbb{R}^m} = \underbrace{\vec{F}(\vec{a})}_{\mathbb{R}^m} + \underbrace{\vec{F}'(\vec{a})}_{m\times m\text{-matrise}}\underbrace{(\vec{x} - \vec{a})}_{\mathbb{R}^m}$

Vi ønsker å løse ligningen $T_{\vec{x}_n}\vec{F}(\vec{x}) = \vec{0}$, dvs

$$\vec{F}(\vec{x}_n) + \vec{F}'(\vec{x}_n)(\vec{x} - \vec{x}_n) = \vec{0}$$

$$\vec{F}'(x_n)^{-1} \;\Big|\; \vec{F}'(\vec{x}_n)(\vec{x} - \vec{x}_n) = -\vec{F}(\vec{x}_n)$$

$$(\vec{x} - \vec{x}_n) = -\vec{F}'(x_n)^{-1}\vec{F}(\vec{x}_n)$$

$$\vec{x} = \vec{x}_n - \vec{F}'(x_n)^{-1}\vec{F}(\vec{x}_n)$$

Vi setter derfor $\vec{x}_{n+1} = \vec{x}_n - \vec{F}'(x_n)^{-1}\vec{F}(\vec{x}_n)$

Newtons metode i flere variable: Tipp (så godt som mulig) en løsning $\vec{x}_0$ og forbedre den ved å bruke iterasjonen

$$\vec{x}_{n+1} = \vec{x}_n - \vec{F}'(x_n)^{-1}\vec{F}(\vec{x}_n),$$

Vil $\vec{x}_n$ konvergere mot en løsning? Ikke alltid, men:

(i) Ja, dersom vi starter tilstrekkelig nær et nullpunkt.

(ii) Kantorovitz' teorem: Dersom det første skrittet $|\vec{x}_1 - \vec{x}_0|$ er lite nok (sammenlignet med $\vec{F}$ og $\vec{F}'$), så konvergerer Newtons metode.

Eksempel: Skal finne nullpunkt(ene) til

$$\vec{F}(x,y) = \begin{pmatrix} x^3 - 3xy^2 + 1 \\ 3x^2y - y^3 \end{pmatrix}$$

ved hjelp av Newtons metode:

$$\boxed{\vec{x}_{n+1} = \vec{x}_n - \vec{F}'(\vec{x}_n)^{-1}\,\vec{F}(\vec{x}_n)}$$

Den deriverte $\vec{F}'$ (dvs Jacobi-matrisen til $\vec{F}$)

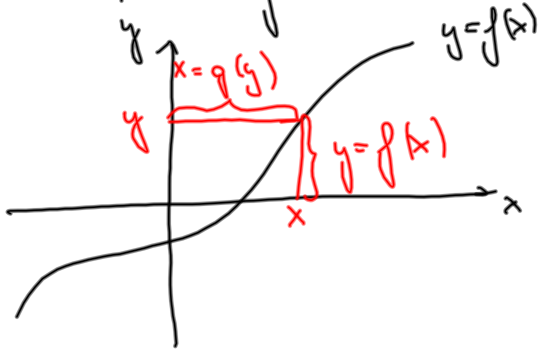
$$\vec{F}'(x,y) = \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 3x^2 - 3y^2 & -6xy \\ 6xy & 3x^2 - 3y^2 \end{pmatrix}$$

I MATLAB: $\vec{u} = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$, $\vec{v} = \vec{F}\begin{pmatrix} x_n \\ y_n \end{pmatrix}$, $A = \vec{F}'(x_n, y_n)$

Løkke: $\vec{u} = \vec{u} - \underbrace{A^{-1}\vec{v}}_{} = \vec{u} - A\backslash\vec{v}$

Omvendte (inverse) funksjoner (5.7)

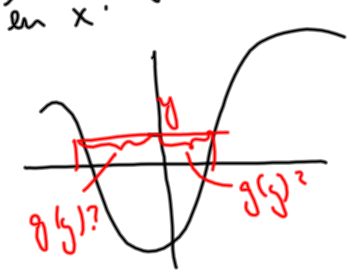
På tallinjen: $f: \mathbb{R} \to \mathbb{R}$



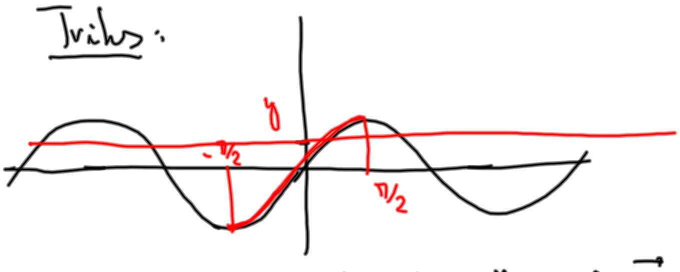
$$g'(y) = \frac{1}{f'(x)} \quad \text{hvor } y = f(x)$$

Forutsetning: f er injektiv; dvs. til hver y hører det høyst én x.

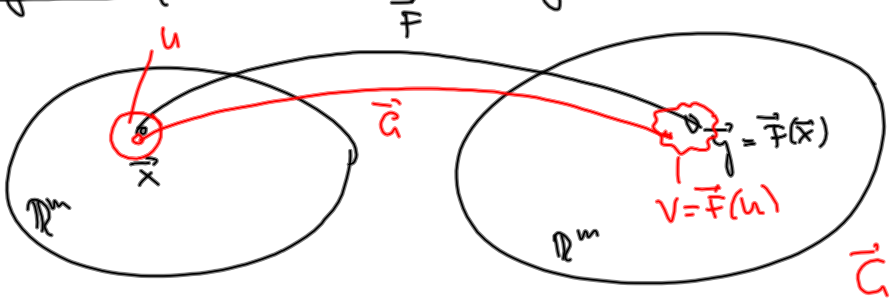
Ikke



Triks:

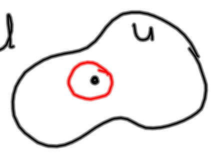


Ny situasjon: $\vec{F}: \mathbb{R}^m \to \mathbb{R}^n$, $\vec{y} = \vec{F}(\vec{x})$



$$\vec{G}(\vec{F}(\vec{z})) = \vec{z}$$

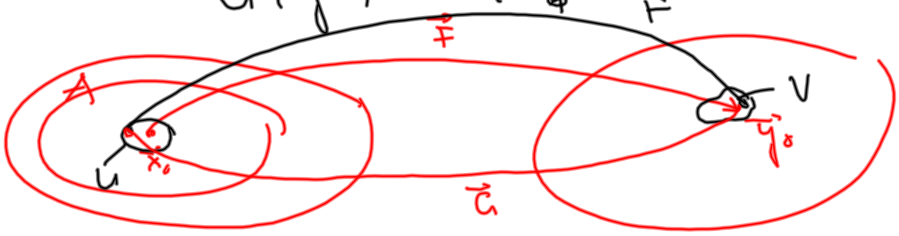
Med en omegn U om $\vec{x}$ mener vi en mengde der $\vec{x}$ er et indre punkt



, dvs at $B(\vec{x}, r) \subset U$ når r er liten nok.

Omvendt funksjonsteorem: Anta at $A \subset \mathbb{R}^m$ og at $\vec{F}: A \to \mathbb{R}^m$ er en deriverbar funksjon og at $\vec{F}'$ er kontinuerlig. Anta videre at $\vec{F}'(\vec{x}_0)$ er inverterbar. Da finnes det en omegn U og $\vec{x}_0$ og en omegn V om $\vec{y}_0 = \vec{F}(\vec{x}_0)$ slik at $V = \vec{F}(U)$ og det finnes en deriverbar funksjon $\vec{G}: V \to U$ slik at $\vec{G}(\vec{y}) = \vec{x} \iff \vec{F}(\vec{x}) = \vec{y}$. Videre

$$\vec{G}'(\vec{y}_0) = \vec{F}'(\vec{x}_0)^{-1}$$



$\vec{G}'(\vec{y}_0) = \vec{F}'(\vec{x}_0)^{-1}$

Hvorfor er $\vec{G}'(\vec{y}_0) = \vec{F}'(\vec{x}_0)^{-1}$?

$$\vec{G}(\vec{F}(\vec{x})) = \vec{x}$$

Deriverer på begge sider:

$$\vec{G}'(\underbrace{\vec{F}(\vec{x})}_{\vec{y}_0})\vec{F}'(\underbrace{\vec{x}}_{\vec{x}_0}) = I_m$$

$$\vec{G}'(y_0)\vec{F}'(x_0) = I_m \Rightarrow \vec{G}'(y_0) = \vec{F}'(x_0)^{-1}$$