4 - 6 - OM DØDSLENGSEL

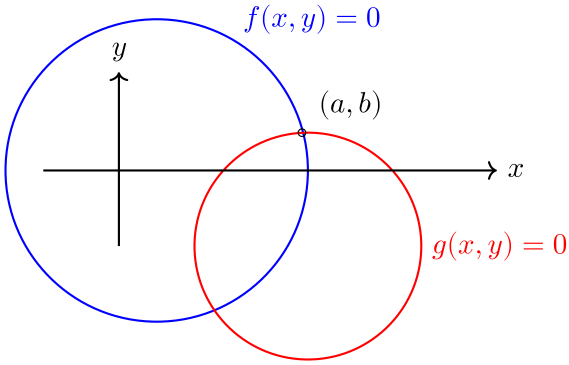
$\boxed{1}$ Anta at vi har to funksjoner $f(x,y)$ og $g(x,y)$. Vi leter etter et punkt (a,b) slik at både

$$f(a,b) = 0 \quad \text{og} \quad g(a,b) = 0.$$

likningene

$$f(x,y) = 0 \quad \text{og} \quad g(x,y) = 0$$

angir nivåkurver for funksjonene f og g. Punktet (a,b) må ligge på skjæringspunktet mellom disse.



Anta at du har en iterasjon (x_n, y_n). Vi setter opp tangentplanene til f og g i (x_n, y_n)

$$z - f(x_n,y_n) = f_x(x_n,y_n)(x - x_n) + f_y(x_n,y_n)(y - y_n)$$

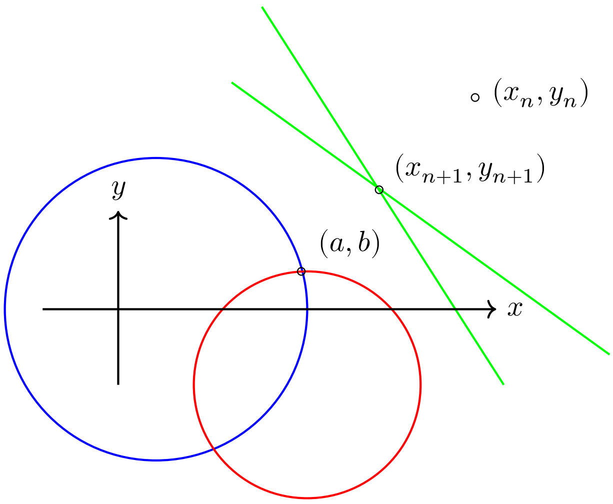
$$z - g(x_n,y_n) = g_x(x_n,y_n)(x - x_n) + g_y(x_n,y_n)(y - y_n).$$

Hvis vi krever at $z = 0$, får vi likninger for skjæringslinjene mellom disse tangentplanene og (x,y)-planet

$$-f(x_n,y_n) = f_x(x_n,y_n)(x - x_n) + f_y(x_n,y_n)(y - y_n)$$

$$-g(x_n,y_n) = g_x(x_n,y_n)(x - x_n) + g_y(x_n,y_n)(y - y_n).$$

Iterasjonen (x_{n+1}, y_{n+1}) defineres som skjæringen mellom disse linjene.



Altså er

$$-f(x_n, y_n) = f_x(x_n, y_n)(x_{n+1} - x_n) + f_y(x_n, y_n)(y_{n+1} - y_n)$$
$$-g(x_n, y_n) = g_x(x_n, y_n)(x_{n+1} - x_n) + g_y(x_n, y_n)(y_{n+1} - y_n),$$

et lineært likningssystem som definerer (x_{n+1}, y_{n+1}). Matrisen er

$$\begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix}$$

og vi skriver

$$-\begin{pmatrix} f(x_n, y_n) \\ g(x_n, y_n) \end{pmatrix} = \begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix} \begin{pmatrix} x_{n+1} - x_n \\ y_{n+1} - y_n \end{pmatrix}.$$

Nå kan vi gange med

$$\begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix}^{-1}$$

fra venstre, legge til (x_n, y_n) på begge sider, og få Newtons metode for systemer

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix}^{-1} \begin{pmatrix} f(x_n, y_n) \\ g(x_n, y_n) \end{pmatrix}.$$

For lesbarhetens skyld lot jeg være å skrive ut hva

$$\begin{pmatrix} f_x(x_n, y_n) & f_y(x_n, y_n) \\ \\ g_x(x_n, y_n) & g_y(x_n, y_n) \end{pmatrix}^{-1}$$

er, men du kan regne det ut ved å huske fra lineæralgebraen at

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Skrevet litt mer kompakt blir dette

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{f}'(\mathbf{x}_n)^{-1}\mathbf{f}(\mathbf{x}_n).$$

$\boxed{2}$ Her er rustkode for denne og oppgave 7 (laga dem i en smell for sammenlikning): `https://folk.ntnu.no/mortano/rust/optimering/`

$\boxed{3}$ Det er jo fordi gradienten er null i et kritisk punkt. Hvis du partiellderiverer

$$V(x, y) = x^3 + x^2y + xy^2 + 2y^3 - x^2 - y^2 - xy - x - y.$$

og setter begge lik null, får du likningsssettet i oppgave 2.

$\boxed{4}$ Formelen

$$\mathbf{x}_{n+1} = \mathbf{x}_n - V''(\mathbf{x}_n)^{-1}V'(\mathbf{x}_n)$$

er den samme som Newtons metode, bare med $V' = \mathbf{f}$.

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \alpha_n \mathbf{p}_n$$

5 Du hiver deg selvfølgelig rett utfor så bratt som overhodet mulig.

6 Når du velger en retning p_n er det sikkert lurt å dra ned til bunnpunktet på kursen du staker ut. Eller er det det? Ikke nødvendigvis, faktisk. Det er ofte like greit å stoppe litt før og finne en ny retning.

7 `https://folk.ntnu.no/mortano/rust/optimering/`

8 Funksjonen

$$V(\mathbf{x}) = \|A\mathbf{x} - \mathbf{b}\|^2$$

er aldri negativ, og måler avstanden mellom $A\mathbf{x}$ og $\mathbf{b}$. Avstanden skal være null om $A\mathbf{x} = \mathbf{b}$ har entydig løsning, og minst mulig ved overbestemte systemer, for eksempel ved regresjon. Dersom du har underbestemte systemer, finnes det gjerne mange $\mathbf{x}$ som gir $V(\mathbf{x}) = 0$.

9 La oss begynne med å derivere $\mathbf{x}^T A\mathbf{x}$. Her må man holde tungen beint i munnen. Det er best å skrive

$$\mathbf{x}^T A\mathbf{x} = \begin{array}{c} a_{11}x_1x_1 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ a_{21}x_2x_1 + a_{12}x_2x_2 + \cdots + a_{2n}x_2x_n \\ \vdots \\ a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_nx_n \end{array}$$

Den deriverte av denne er rekkevektoren

$$\begin{aligned} \nabla \mathbf{x}^T A\mathbf{x} &= \begin{array}{c} (2a_{11}x_1 + (a_{12} + a_{21})\,x_2 + \cdots + (a_{1n} + a_{n1})\,x_n, \\ (a_{21} + a_{12})\,x_1 + 2a_{22}x_2 + \cdots + (a_{2n} + a_{n2})\,x_n, \\ \vdots \\ (a_{n1} + a_{1n})\,x_1 + (a_{n2} + a_{2n})\,x_2 + \cdots + 2a_{nn}x_n) \end{array} \\ &= \mathbf{x}^T\left(A + A^T\right) \\ &= 2\mathbf{x}^T A \end{aligned}$$

Dersom $\mathbf{x}$ er en søylevektor, får vi altså

$$\nabla V(\mathbf{x}) = \mathbf{x}^T A - \mathbf{b}^T = (A\mathbf{x} - \mathbf{b})^T$$

som har entydig nullpunkt i $A^{-1}\mathbf{b}$ siden A er positivt definitt, og dermed er inverterbar siden ingen av egenverdiene kan være null. Hessematrisen blir bare A, og siden denne er positivt definitt må det kritiske punktet være et minimumspunkt.

10 La

$$V(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{b}.$$

Vi bruker kjerneregelen og setter lik null:

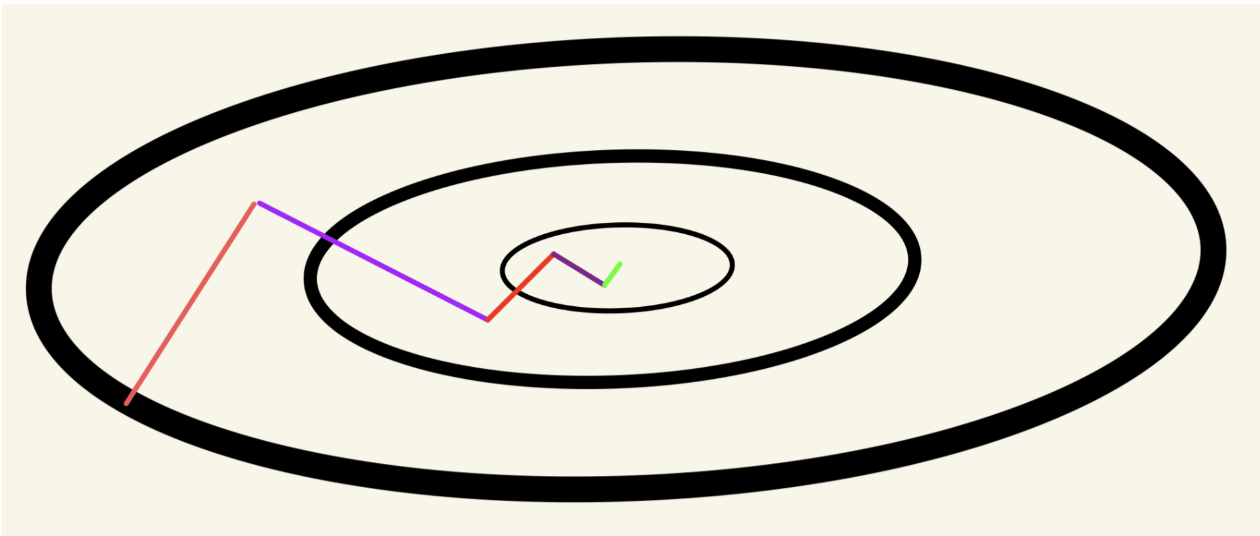
$$\begin{aligned}\frac{d}{d\alpha_n}V(\mathbf{x}_n+\alpha_n\mathbf{p}_n) &= \nabla V(\mathbf{x}_n+\alpha_n\mathbf{p}_n)^T\mathbf{p}_n \\ &= \left(A\left(\mathbf{x}_n+\alpha_n\mathbf{p}_n\right)-\mathbf{b}\right)^T\mathbf{p}_n = 0\end{aligned}$$

som gir

$$\alpha_n = \frac{\left(\mathbf{b}-A\mathbf{x}_n\right)^T\mathbf{p}_n}{\mathbf{p}_n^T A\mathbf{p}_n}.$$

Dette kan gjøres for generelle andreordens polynomer, men vi skal bare studere denne, så jeg plager deg ikke med det generelle tilfellet.

11



12 Her er det relativt enkelt å modifisere koden i oppgave 2.

13 Dette var pensum i TMA4106 og jeg tror vi sier oss ferdige med det.

14 Vi ganger likningen

$$\mathbf{x} = \sum_k \alpha_k\mathbf{p}_k.$$

med $\mathbf{p}_l^T A$ fra venstre og får

$$\mathbf{p}_l^T A\mathbf{x} = \sum_k \alpha_k\mathbf{p}_l^T A\mathbf{p}_k = \alpha_k\mathbf{p}_l^T A\mathbf{p}_k$$

og bruker at $A\mathbf{x}=\mathbf{b}$, slik at

$$\alpha_k = \frac{\mathbf{p}_k^T\mathbf{b}}{\mathbf{p}_k^T A\mathbf{p}_k} = \frac{(\mathbf{p}_k,\mathbf{b})}{(\mathbf{p}_k,\mathbf{p}_k)_A}.$$

$\boxed{15}$ Samme triks, vi ganger med $\mathbf{p}_k^T A$ fra venstre og krever at indreproduktet skal bli null:

$$
\begin{aligned}
0 =& \mathbf{p}_k^T A \mathbf{p}_{k+1} \\
=& \mathbf{p}_k^T A \left(-\mathbf{r}_{k+1} + \beta_{k+1} \mathbf{p}_k\right) \\
=& -\mathbf{p}_k^T A \mathbf{r}_{k+1} + \beta_{k+1} \mathbf{p}_k^T A \mathbf{p}_k
\end{aligned}
$$

slik at

$$
\beta_{k+1} = \frac{(\mathbf{r}_{k+1}, \mathbf{p}_k)_A}{(\mathbf{p}_k, \mathbf{p}_k)_A}.
$$

$\boxed{16}$ Dette er litt hardt for TMA4121, ihvertfall i år. Kanskje neste år.

$\boxed{17}$ Dette er for hardt for eksamen.