TMA4135 Matematikk 4D

Forelesningsnotater

Ligningsløsere

Disse forelesningsnotatene er ikke noen erstatning for læreboken, for den er en murstein, altså tar den flere tiår, en haug med slaver, samt et stort forlag, å skrive. Feil sniker seg inn - finner du noen, send gjerne epost til morten.nome@gmail.com.

Ligningsløsere for envariable ligninger

I anvendelser møter man ligninger som ikke lar seg løse analytisk. Et klassisk eksempel er $x = \cos x$.

Fikspunktiterasjonen

Anta at vi har en ligning på formen

$$r = g(r).$$

Iterasjonen

$$x_{i+1} = g(x_i). \tag{1}$$

vil konvergere mot r dersom $r, x_0 \in (a, b)$ og $g'(x) < 1$ på (a, b).

Newtons metode

Anta at ligningen er på formen

$$f(r) = 0.$$

Prosedyren for Newtons metode er som følger: slå tangenten til funksjonen f i iterasjonen x_i. Punktet der tangenten skjærer x-aksen er den nye iterasjonen x_{i+1}. Hvis vi setter opp ligningen for tangenten

$$y - f(x_i) = f'(x_i)(x - x_i),$$

og sier at iterasjonen x_{i+1} er de x-verdien slik at $y = 0$

$$-f(x_i) = f'(x_i)(x_{i+1} - x_i).$$

blir iterasjonen

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}.$$

Sekantmetoden

Denne ligner på Newton, men vi slår sekanten til f gjennom iterasjonene x_i og x_{i-1}, istedet for tangenten i x_i. Punktet der sekanten skjærer x-aksen er den nye iterasjonen x_{i+1}. Iterasjonen blir

$$x_{i+1} = x_i - f(x_i)\frac{x_i - x_{i-1}}{f(x_i) - f(x_{i-1})}.$$

Merk at her må man ha to initialgjetninger x_1 og x_0, siden man trenger to punkter for å slå en sekant.

Litt om konvergens

Vi bruker taylorutvikling for å analysere hvor fort en ligningsløser konvergerer mot løsningen. La oss begynne med fikspunktiterasjonen. Vi taylorutvikler g om r, og får

$$g(x_n) = g(r) + g'(r)(x_n - r) + \frac{g''(r)}{2}(x_n - r)^2 + \frac{g'''(r)}{3!}(x_n - r)^3 + \dots$$

Nå kan vi skrive

$$x_{n+1} - r = g(x_n) - g(r) = g'(r)(x_n - r) + \frac{g''(r)}{2}(x_n - r)^2 + \frac{g'''(r)}{3!}(x_n - r)^3 + \dots$$

og da har vi en ligning som sier noe om størrelsen på $x_{n+1} - r$ som en funksjon av $x_n - r$, altså hvor mye feilen minker fra iterasjon til iterasjon. For fikspunktiterasjonen ser vi at det blir best konvergens om g' er så liten som mulig. Dersom $g' < 1$ på et intervall som inneholder x_0 og r, kan konvergens garanteres. Konvergensen kalles for øvrig lineær, siden $x_{n+1} - r$ er proporsjonal med $x_n - r$.

For å analysere Newtons metode

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)},$$

må vi se på det som en fikspunktiterasjon der

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Vi beregner

$$g'(x) = 1 - \frac{\left(f(x)\right)^2 - f(x)f''(x)}{\left(f(x)\right)^2} = -\frac{f(x)f''(x)}{\left(f(x)\right)^2}$$

Husk at $f(r) = 0$. Dersom $f'(r) \neq 0$, får vi

$$g'(r) = -\frac{f(r)f''(r)}{\left(f(r)\right)^2} = 0,$$

slik at

$$x_{n+1} - r = g(x_n) - g(r) = \frac{g''(r)}{2}(x_n - r)^2 + O((x_n - r)^3).$$

Dette kalles kvadratisk konvergens - feilen etter $n+1$ iterasjoner er proporsjonal med kvadratet av feilen etter n iterasjoner. Da konvergerer det fort.

Når det gjelder sekantmetoden, kan det vises at

$$x_{n+1} - r = C(x_n - r)^{1.62},$$

men det er et svineri, du må taylorutvikle noe ut av det hinsidige. Konvergens kalles superlineær, altså kjappere enn fikspunkt, men treigere enn Newton.

Ligningsløsere for ligningssystem

Vi skal ta for oss to metoder for å løse ligningsystemer - en fikspunktiterasjon for lineære systemer, og Newtons metode for ikkelineære systemer.

Fikspunktiterasjon for lineære ligningssystemer

Vi skriver om ligningssystemet

$$Ax = b \tag{2}$$

til formen

$$x = G(x).$$

Vi skriver nå

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

definerer

$$D = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

og

$$R = \begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{pmatrix},$$

og setter opp iterasjonen

$$x_{n+1} = D^{-1}(b - Rx_n).$$

Dette kalles Jacobis metode. Det finnes en variant som konvergerer litt kjappere. Vi skriver opp Jacobi

$$\begin{aligned} x_{n+1} &= (b_1 - a_{12}y_n - a_{13}z_n)/a_{11} \\ y_{n+1} &= (b_2 - a_{21}x_n - a_{23}z_n)/a_{22} \\ z_{n+1} &= (b_2 - a_{31}x_n - a_{32}y_n)/a_{33}. \end{aligned}$$

Nå beregner man x_{n+1} først. Når man så skal beregne y_{n+1} er det ingenting i veien for å bruke x_{n+1} istedet for x_n, og når man skal beregne z_{n+1}, er det ingenting i veien for å bruke x_{n+1} og y_{n+1} istedet for x_n og y_n. Metoden blir

$$\begin{aligned} x_{n+1} &= (b_1 - a_{12}y_n - a_{13}z_n)/a_{11} \\ y_{n+1} &= (b_2 - a_{21}x_{n+1} - a_{23}z_n)/a_{22} \\ z_{n+1} &= (b_2 - a_{31}x_{n+1} - a_{32}y_{n+1})/a_{33}, \end{aligned}$$

og heter Gauss-Seidels metode.

Man kan få en magefølelse for konvergensen som følger. Dersom $Ax = b$, vil

$$x = D^{-1}(b - Rx),$$

siden dette bare er en omskriving av $Ax = b$. Vi kan trekke iterasjonsligningen fra den eksakte ligningen og få

$$x_{n+1} - x = D^{-1}R(x_n - x).$$

Her ser vi at feilen minker dersom matrisen $D^{-1}R$ har en 'krympende virkning' på vektorer du ganger inn i den. Vi har

$$D^{-1} = \begin{pmatrix} 1/a_{11} & 0 & 0 \\ 0 & 1/a_{22} & 0 \\ 0 & 0 & 1/a_{33} \end{pmatrix},$$

og følgelig

$$D^{-1}R = \begin{pmatrix} 0 & a_{12}/a_{11} & a_{13}/a_{11} \\ a_{21}/a_{22} & 0 & a_{23}/a_{22} \\ a_{31}/a_{33} & a_{32}/a_{33} & 0 \end{pmatrix},$$

Hvis matrisen $D^{-1}R$ skal ha en 'krympende virkning' på vektorer som ganges inn i den, bør komponentene i $D^{-1}R$ være minst mulig. Vi ser nå at dette skjer dersom diagonalelementene i A er mye større en de andre elementene i A. Jo større forskjell, desto raskere konvergens. Man kan vise at Jacobis iterasjon konvergerer dersom vi har såkalt diagonaldominans, nemlig at

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$$

for alle rekker i A.

Newtons metode i to dimensjoner

Anta at vi har to funksjoner $f_1(x,y)$ og $f_2(x,y)$. Vi leter etter et punkt (a,b) slik at både

$$f_1(a,b) = 0$$

og

$$f_2(a,b) = 0.$$

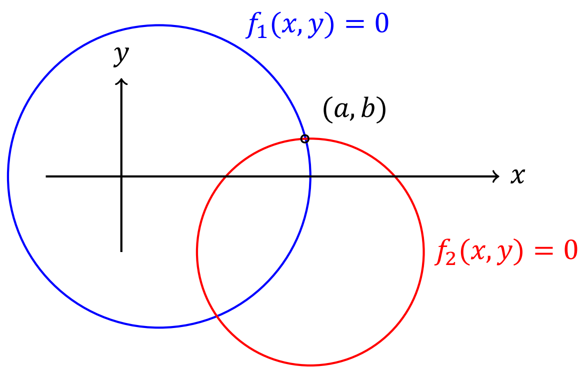
Ligningene

$$f_1(x,y) = 0$$

og

$$f_2(x,y) = 0$$

angir nivåkurver for funksjonene f_1 og f_2. 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_1 og f_2 i (x_n, y_n)

$$z - f_1(x_n, y_n) = \frac{\partial f_1}{\partial x}(x_n, y_n)(x - x_n) + \frac{\partial f_1}{\partial y}(x_n, y_n)(y - y_n)$$

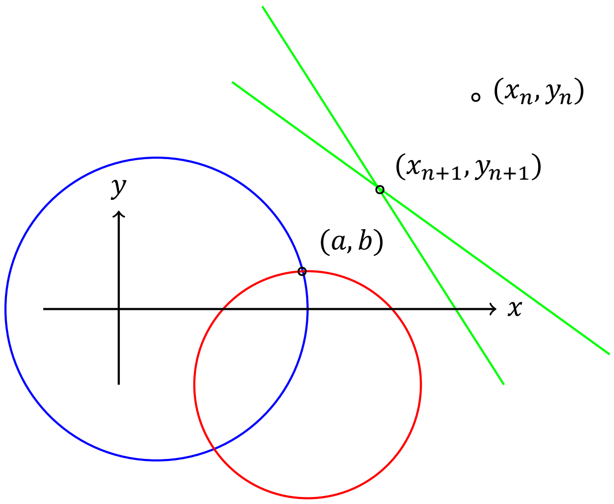
$$z - f_2(x_n, y_n) = \frac{\partial f_2}{\partial x}(x_n, y_n)(x - x_n) + \frac{\partial f_2}{\partial y}(x_n, y_n)(y - y_n).$$

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

$$-f_1(x_n, y_n) = \frac{\partial f_1}{\partial x}(x_n, y_n)(x - x_n) + \frac{\partial f_1}{\partial y}(x_n, y_n)(y - y_n)$$

$$-f_2(x_n, y_n) = \frac{\partial f_2}{\partial x}(x_n, y_n)(x - x_n) + \frac{\partial f_2}{\partial y}(x_n, y_n)(y - y_n).$$

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



Altså er

$$-f_1(x_n, y_n) = \frac{\partial f_1}{\partial x}(x_n, y_n)(x_{n+1} - x_n) + \frac{\partial f_1}{\partial y}(x_n, y_n)(y_{n+1} - y_n)$$

$$-f_2(x_n, y_n) = \frac{\partial f_2}{\partial x}(x_n, y_n)(x_{n+1} - x_n) + \frac{\partial f_2}{\partial y}(x_n, y_n)(y_{n+1} - y_n).$$

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

$$\begin{pmatrix} \frac{\partial f_1}{\partial x}(x_n, y_n) & \frac{\partial f_1}{\partial y}(x_n, y_n) \\ \frac{\partial f_2}{\partial x}(x_n, y_n) & \frac{\partial f_2}{\partial y}(x_n, y_n) \end{pmatrix}$$

og vi skriver

$$-\begin{pmatrix} f_1(x_n, y_n) \\ f_2(x_n, y_n) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x_n, y_n) & \frac{\partial f_1}{\partial y}(x_n, y_n) \\ \frac{\partial f_2}{\partial x}(x_n, y_n) & \frac{\partial f_2}{\partial y}(x_n, y_n) \end{pmatrix} \begin{pmatrix} x_{n+1} - x_n \\ y_{n+1} - y_n \end{pmatrix}.$$

Nå kan vi gange med

$$\begin{pmatrix} \frac{\partial f_1}{\partial x}(x_n, y_n) & \frac{\partial f_1}{\partial y}(x_n, y_n) \\ \frac{\partial f_2}{\partial x}(x_n, y_n) & \frac{\partial f_2}{\partial y}(x_n, y_n) \end{pmatrix}^{-1}$$

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

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} \frac{\partial f_1}{\partial x}(x_n, y_n) & \frac{\partial f_1}{\partial y}(x_n, y_n) \\ \frac{\partial f_2}{\partial x}(x_n, y_n) & \frac{\partial f_2}{\partial y}(x_n, y_n) \end{pmatrix}^{-1} \begin{pmatrix} f_1(x_n, y_n) \\ f_2(x_n, y_n) \end{pmatrix}.$$

Jeg gadd ikke skrive ut hva

$$\begin{pmatrix} \frac{\partial f_1}{\partial x}(x_n, y_n) & \frac{\partial f_1}{\partial y}(x_n, y_n) \\ \frac{\partial f_2}{\partial x}(x_n, y_n) & \frac{\partial f_2}{\partial y}(x_n, y_n) \end{pmatrix}^{-1}$$

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

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