

Newton's metod i flera variabler

Vi har tidigare lärt oss hur man kan använda Newtons metod för att lösa olinjära ekvationer i en variabel med numeriska beräkningar. Vi kommer nu lära oss en metod som är en generalisering av Newtons metod, för flera variabler. Vi har n obekanta variabler och n ekvationer som vi sätter ihop på vektorform, så att vi får denna problemformulering:

Problem 1.0.1. Givet $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^n$, hitta $\mathbf{x}_* \in \mathbb{R}^n$ så att

$$\mathbf{f}(\mathbf{x}_*) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n.$$

Partiella derivator och Jacobimatrisen

I Newtons metod (i en variabel) behöver vi derivatan av funktionen. När vi vill härleda Newtons metod i flera variabler behöver vi därför lära oss det naturliga sättet att derivera i flera variabler. Vi använder oss av så kallade partiella derivator. För enkelhetens skull betraktar vi först en reell funktion i två variabler.

Om $g = g(x, y)$ är en reellvärd funktion av två variabler så definieras de partiella derivatorna av g i punkten (a, b) i definitionsmängden som

$$\frac{\partial g}{\partial x}(a, b) := \lim_{h \rightarrow 0} \frac{g(a+h, b) - g(a, b)}{h} \quad (1.1)$$

$$\frac{\partial g}{\partial y}(a, b) := \lim_{h \rightarrow 0} \frac{g(a, b+h) - g(a, b)}{h} \quad (1.2)$$

En partiell derivata av en funktion i flera variabler motsvarar alltså derivatan i en variabel, när de andra variablerna hålls konstanta.

Innan du läser detta är det bra om du läst i kursboken Sauer kap 1.4

Partiella derivator och Jacobimatrisen kan man lära sig mer om i kursen SF1626 som är valfri för CDATE i högre årskurser. I den här kursen lär vi oss endast det nödvändiga för att härleda och använda Newtons metod i flera variabler.

 Exempel partiell derivata

Om $g(x, y) = x^2 + 3xy$ är de partiella derivatorna

$$\begin{aligned}\frac{\partial g}{\partial x}(x, y) &:= 2x + 3y \\ \frac{\partial g}{\partial y}(x, y) &:= 3x\end{aligned}$$



I vårt sammanhang, dvs när $\mathbf{f}$ är en funktion från $\mathbb{R}^n$ till $\mathbb{R}^n$, tolkar i partiella derivator kolumnvis. Om vi kallar elementen i $\mathbf{f}$ för $f_1, \dots, f_n$, dvs

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{bmatrix}$$

så är den partiella derivatan vektorn

$$\frac{\partial \mathbf{f}}{\partial x_1} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_1, \dots, x_n) \\ \vdots \\ \frac{\partial f_n}{\partial x_1}(x_1, \dots, x_n) \end{bmatrix}.$$

Vi behöver ytterligare en definition. Vi kallar den matris som innehåller alla partiella derivator för Jacobimatrisen.

Definition 1.1.1. Givet en funktion $\mathbf{f}(\mathbf{x})$, definierar vi Jacobimatrisen som

$$\begin{aligned}J(\mathbf{x}) &= \left[\frac{\partial \mathbf{f}}{\partial x_1}, \dots, \frac{\partial \mathbf{f}}{\partial x_n} \right] \\ &= \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}\end{aligned}$$

 Exempel Jacobimatris

Funktionen $\mathbf{f}$ given som

$$\mathbf{f}(x_1, x_2) = \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix} = \begin{bmatrix} x_1^2 + 3x_1x_2 \\ \sin(x_1x_2) + x_1 \end{bmatrix}$$

har Jacobimatris

$$J(x_1, x_2) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 & 3x_1 \\ x_2 \cos(x_1x_2) + 1 & x_1 \cos(x_1x_2) \end{bmatrix}$$



Härledning av Newtons metod i flera variabler

Anledningen till att vi introducerade Jacobimatriser ovan är att de tillåter oss att göra en Taylorutveckling i flera variabler. Jacobimatrisen är en matris som generaliserar derivatabegreppet till flera variabler. För funktioner som är tillräckligt ofta deriverbara kan vi utveckla:

$$\mathbf{f}(\mathbf{x} + \mathbf{h}) = \mathbf{f}(\mathbf{x}) + J(\mathbf{x})\mathbf{h} + \mathcal{O}(\|\mathbf{h}\|^2). \quad (1.3)$$

Med hjälp av denna utveckling kan vi härleda Newtons metod i flera variabler på samma sätt som vi härledde Newtons metod i en variabel.

Låt oss anta att $\mathbf{x}_0 \in \mathbb{R}^n$ är givet (och kommer att tolkas som en approximation till lösningen $\mathbf{x}_*$). Som en direkt konsekvens av (1.3), gäller

$$\mathbf{0} = \mathbf{f}(\mathbf{x}_*) = \mathbf{f}(\mathbf{x}_0 + \mathbf{x}_* - \mathbf{x}_0) = \mathbf{f}(\mathbf{x}_0) + J(\mathbf{x}_0)(\mathbf{x}_* - \mathbf{x}_0) + \mathcal{O}(\|\mathbf{x}_* - \mathbf{x}_0\|^2). \quad (1.4)$$

Utifrån denna approximation definierar vi nu en ny approximation $\mathbf{x}_1$, som förhoppningsvis är bättre än $\mathbf{x}_0$. Mer exakt, *definierar* vi $\mathbf{x}_1$ som den vektor som uppfyller ekvation (1.4) när vi tagit bort (försummat) den kvadratiske termen och ersatt $\mathbf{x}_*$ med $\mathbf{x}_1$:

$$\mathbf{0} = \mathbf{f}(\mathbf{x}_0) + J(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_0). \quad (1.5)$$

Eftersom vi definierar $\mathbf{x}_1$ genom att försumma den kvadratiske termen, förväntar vi oss att $\mathbf{x}_1$ är en mycket bra approximation $\mathbf{x}_*$ om $\|\mathbf{x}_0 - \mathbf{x}_*\|^2$ är litet, dvs om $\mathbf{x}_0$ är en approximation av $\mathbf{x}_*$. Under antagandet att matrisen $J(\mathbf{x}_0)$ är inverterbar kan ekvation (1.5) skrivas om som

$$\mathbf{x}_1 = \mathbf{x}_0 - J(\mathbf{x}_0)^{-1}\mathbf{f}(\mathbf{x}_0)$$

Vi kan alltså beräkna $\mathbf{x}_1$ givet $\mathbf{x}_0$. Detta resonemang kan upprepas och motiverar definitionen av Newtons metod.

Definition 1.2.1 (Newtons metod). Givet $\mathbf{x}_0$ och funktion $\mathbf{f}(\mathbf{x})$ med Jacobimatriss $J(\mathbf{x})$, kallas iterationen

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mathbf{h}_i, \text{ med } i = 1, 2, \dots,$$

och

$$\mathbf{h}_i = J(\mathbf{x}_i)^{-1}\mathbf{f}(\mathbf{x}_i),$$

för Newtons metod i flera variabler.

Exempel Newton flera variabler

Låt oss säga att vi vill lösa ekvationssystemet

$$\begin{aligned} 0 &= x_1 + x_2^2 \\ 1 &= x_1^2 + 4x_2 \end{aligned}$$

Newtons metod i flera variabler finns även beskrivet i kursboken Sauer kapitel 2.7.1.

Observera att $\mathbf{f}(\mathbf{x})$, $\mathbf{x}$ och $\mathbf{h}$ är vektorer av längd n i ekvation (1.3) och Jacobimatrissen $J(\mathbf{x})$ är en $n \times n$ -matrix

Definition 1.2.1 generaliserar Newtons metod i en variabel. Se kursboken Sauer kapitel 1.4.

Vi sätter upp vår ekvation på formen i Problem 1.0.1:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1 + x_2^2 \\ -1 + x_1^2 + 4x_2 \end{bmatrix}$$

och får Jacobimatrisen

$$J(\mathbf{x}) = \begin{bmatrix} 1 & 2x_2 \\ 2x_1 & 4 \end{bmatrix}.$$

Vi behöver en startgissning för att genomföra Newtons metod. Vi ser direkt att den första ekvationen är uppfylld om $x_1 = x_2 = 0$, och provar först med det som startgissning. Eftersom

$$f\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \text{ och } J\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

blir

$$\mathbf{x}_1 = \mathbf{x}_0 - J(\mathbf{x}_0)^{-1}\mathbf{f}(\mathbf{x}_0) = \begin{bmatrix} 0 \\ -1/4 \end{bmatrix}.$$

Upprepning kan göras i MATLAB på följande sätt. I programmet nedan stannar iterationen när *korrektionen* är $\mathbf{h}$ är tillräckligt liten:

```
>> f=@(x) [x(1)+x(2)^2; -1+x(1)^2+4*x(2)];
>> J=@(x) [1, 2*x(2); 2*x(1), 4];
>> h=inf;
>> TOL=1e-10;
>> x=[0;0];
>> while (norm(h)>TOL)
>>     h=J(x)\f(x);
>>     x=x-h;
>> end
>> h
h =
    1.0e-12 *
    -0.0883
     0.1918
>> x =
    -0.0620
     0.2490
>> f(x)
ans =
    1.0e-17 *
     0.6939
         0
```

Precis som för Newtons metod i en variabel finns många olika sätt att välja startgissningar, och det finns inget allmänt tillvägagångssätt. I praktiken försöker man ofta använda tillämpningspecifika egenskaper för att få fram en startgissning. Man kör sedan metoden och se om det fungerar. Om den inte fungerar får man gå tillbaka och försöka förbättra startgissningen.

