



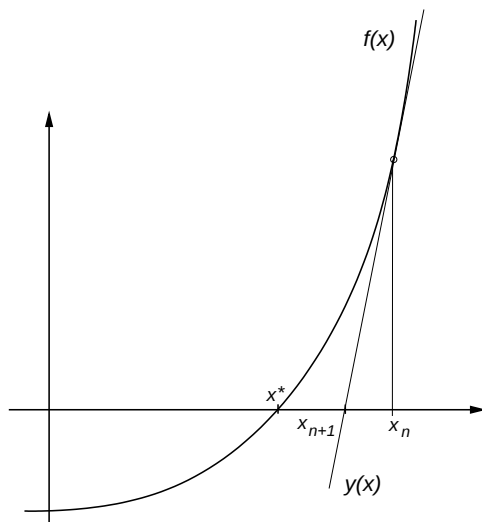
KTH Computer Science
and Communication

Övning 1

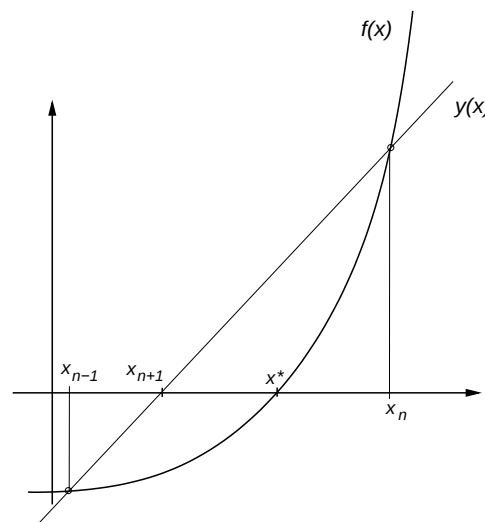
Konvergensordning för sekantmetoden

Sekantmetoden är en förenklad version av Newtons metod för att lösa skalära olinjära ekvationer $f(x) = 0$. Till skillnad från Newtons metod använder sig sekantmetoden inte av derivatan till f , vilken kan vara svår att beräkna när f är komplicerad. I båda metoderna approximeras f med en enkel funktion i varje iteration, och nästa tal x_{n+1} ges som lösningen till den förenklade ekvationen. I Newtons metod (figur a) används tangenten till f i x_n , som approximation av f för att beräkna x_{n+1} ,

$$y(x) = f(x_n) + (x - x_n)f'(x_n), \quad y(x_{n+1}) = 0 \Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$



(a) Newtons metod



(b) Sekantmetoden

Sekantmetoden (figur b) använder istället den linje som går genom $f(x_n)$ och $f(x_{n-1})$,

$$y(x) = f(x_n) + (x - x_n) \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

Nästa approximation x_{n+1} definieras även här som lösningen till den förenklade ekvationen, $y(x_{n+1}) = 0$, vilket ger formeln för sekantmetoden:

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

Kommentarer:

- Sekantmetoden behöver spara de två senaste approximationerna, x_{n-1} och x_n (inte bara x_n som för Newton).
- Ingen derivata behöver räknas ut. Den approximeras av kvoten $(f(x_n) - f(x_{n-1})) / (x_n - x_{n-1})$.
- Kan generaliseras till system, men det är komplicerat.

Konvergensordning

Låt den exakta roten till f vara x^* och definiera e_n som felet i iteration n ,

$$f(x^*) = 0, \quad e_n := x_n - x^*.$$

En konvergent sekvens har konvergensordning p om kvoten $|e_{n+1}|/|e_n|^p$ konvergerar mot ett nollskilt tal, dvs om

$$\lim_{n \rightarrow \infty} x_n \rightarrow x^* \quad \Rightarrow \quad \lim_{n \rightarrow \infty} \frac{|e_{n+1}|}{|e_n|^p} \rightarrow C \neq 0.$$

Vi undersöker nu vilken konvergensordning sekvenser som genereras med sekantmetoden har. Vi antar då att andraderivatan vid roten är nollskild, $f''(x^*) \neq 0$. Efter subtraktion med x^* i (1) får vi

$$\begin{aligned} e_{n+1} &= e_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})} = \frac{e_n[f(x_n) - f(x_{n-1})] - f(x_n)(e_n - e_{n-1})}{f(x_n) - f(x_{n-1})} \\ &= \frac{e_{n-1}f(x_n) - e_nf(x_{n-1})}{f(x_n) - f(x_{n-1})} = e_ne_{n-1} \frac{\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}}{\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}} =: e_ne_{n-1}g_n. \end{aligned}$$

Låt $s(x) := f(x)/(x - x^*)$ och $\Delta x_n = x_n - x_{n-1} = e_n - e_{n-1}$. Det ger att

$$g_n = \frac{s(x_{n-1} + \Delta x_n) - s(x_{n-1})}{f(x_{n-1} + \Delta x_n) - f(x_{n-1})} = \frac{s'(x_{n-1}) + O(\Delta x_n)}{f'(x_{n-1}) + O(\Delta x_n)}.$$

Vi undersöker derivatan av s noggrannare,

$$\begin{aligned} s'(x_{n-1}) &= -\frac{f(x_{n-1}) + (x^* - x_{n-1})f'(x_{n-1})}{(x_{n-1} - x^*)^2} = -\frac{f(x^*) - \frac{1}{2}f''(x_{n-1})e_{n-1}^2 + O(e_{n-1}^3)}{e_{n-1}^2} \\ &= \frac{1}{2}f''(x_{n-1}) + O(e_{n-1}). \end{aligned}$$

När metoden konvergerar går e_n , och således också $|\Delta x_n|$, mot noll, dvs så snart $x_n \rightarrow x^*$ följer att

$$\frac{e_{n+1}}{e_ne_{n-1}} = g_n = \frac{\frac{1}{2}f''(x_{n-1}) + O(|\Delta x_n| + |e_{n-1}|)}{f'(x_{n-1}) + O(\Delta x_n)} \rightarrow \frac{f''(x^*)}{2f'(x^*)} \neq 0. \quad (2)$$

Detta är inte riktigt på den form som definitionen av konvergensordning använder, eftersom kvoten innehåller både e_n och e_{n-1} . Vi behöver därför undersöka vad konsekvensen är för andra kvoter $|e_{n+1}|/|e_n|^p$. Vi börjar med att definiera vektorerna

$$w_n := \begin{pmatrix} \log |e_n| \\ \log |e_{n-1}| \end{pmatrix}, \quad d_n = \begin{pmatrix} \log |g_n| \\ 0 \end{pmatrix}.$$

Efter logaritmering ger (2) att

$$\log |e_{n+1}| = \log |e_n| + \log |e_{n-1}| + \log |g_n|.$$

Vi har därför sambandet

$$w_{n+1} = Bw_n + d_n, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Matrisen B är symmetrisk och enligt resultat i linjär algebra kan den då *diagonaliseras* med ortonormala matriser, dvs vi kan skriva den på formen

$$B = S\Lambda S^T, \quad S = \begin{pmatrix} | & | \\ r_1 & r_2 \\ | & | \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad S^T S = I,$$

där r_j är normerade egenvektorer till B och λ_j är motsvarande egenvärden,

$$\lambda_1 = \frac{1}{2}(1 - \sqrt{5}), \quad \lambda_2 = \frac{1}{2}(1 + \sqrt{5}), \quad r_1 = \frac{|\lambda_1|}{\sqrt{1 + \lambda_1^2}} \begin{pmatrix} 1 \\ 1/\lambda_1 \end{pmatrix}, \quad r_2 = \frac{|\lambda_2|}{\sqrt{1 + \lambda_2^2}} \begin{pmatrix} 1 \\ 1/\lambda_2 \end{pmatrix}.$$

Vi multiplicerar med S^T och får

$$S^T w_{n+1} = S^T S \Lambda S^T w_n + S^T d_n \quad \Rightarrow \quad v_{n+1} = \Lambda v_n + S^T d_n, \quad v_n := S^T w_n.$$

Eftersom Λ är diagonal kopplar elementen i $v_n = (v_{1,n} \ v_{2,n})^T$ isär. Vi betraktar första komponenten.

$$v_{1,n+1} = \lambda_1 v_{1,n} + r_{1,1} \log g_n.$$

Denna rekursion har lösningen

$$v_{1,n+1} = v_{1,0} \lambda_1^{n+1} + r_{1,1} \sum_{j=0}^n \lambda_1^{n-j} \log |g_j|.$$

Eftersom $|\lambda_1| < 1$ går första termen mot noll när $n \rightarrow \infty$. Vidare, är sekvensen $\{\log |g_n|\}$ begränsad eftersom $\{|g_n|\}$ konvergerar mot ett nollskilt tal. Det betyder att det finns ett tal D så att

$$\sum_{j=0}^n |\lambda_1^{n-j} \log g_j| \leq D \sum_{j=0}^n |\lambda_1^{n-j}| = D \sum_{j=0}^n |\lambda_1^j| \leq D \sum_{j=0}^{\infty} |\lambda_1|^j = \frac{D}{1 - |\lambda_1|} < \infty.$$

Summan är alltså absolutkonvergent och därmed konvergent mot ett tal, säg q . Vi får

$$\lim_{n \rightarrow \infty} v_{1,n} = r_{1,1} q.$$

Men

$$v_{1,n+1} = (S^T w_{n+1})_1 = r_1^T w_{n+1} = r_{1,1} \left(\log |e_{n+1}| + \frac{1}{\lambda_1} \log |e_n| \right).$$

Därför får vi slutligen

$$\frac{|e_{n+1}|}{|e_n|^{-1/\lambda_1}} = \exp \left(\frac{v_{1,n+1}}{r_{1,1}} \right) \rightarrow \exp(q) \neq 0.$$

Följdaktligen är konvergensordningen $p = -1/\lambda_1 = (1 + \sqrt{5})/2 \approx 1.62$. Detta är något långsammare än Newton ($p = 2$) men snabbare än fixpunktiterationer ($p = 1$).