ALA-a 2005

RÄKNEÖVNING — VECKA 7

David Heintz, 14 oktober 2005

Innehåll

1 Lite teori

1.1 Kapitel 7

Visa att en periodisk decimalutveckling är ett rationellt tal.

Vi skriver en oändlig n-periodisk decimalutveckling som

$$p = .q_1q_2\cdots q_nq_1q_2\cdots q_n\cdots\,, \tag{1}$$

där varje period utgörs av siffrorna $q_1q_2\cdots q_n$. Trunkeras (1) efter m perioder fås

$$p_m = \frac{q_1q_2\cdots q_n}{10^n} + \frac{q_1q_2\cdots q_n}{10^{2n}} + \cdots + \frac{q_1q_2\cdots q_n}{10^{mn}}, \tag{2}$$

med varje term motsvarande n decimaler i utvecklingen. Faktorisera (2) och utnyttja att vi har en geometrisk summa

$$\begin{aligned} p_m &= \frac{q_1q_2\cdots q_n}{10^n}\left(1 + \frac{q_1q_2\cdots q_n}{10^n} + \cdots + \frac{q_1q_2\cdots q_n}{10^{(m-1)n}}\right) \\ &= \frac{q_1q_2\cdots q_n}{10^n}\left(\frac{1-10^{-mn}}{1-10^{-n}}\right) = \frac{q_1q_2\cdots q_n}{10^n}\frac{10^n}{10^n-1}\left(1-10^{-mn}\right) \\ &= \frac{q_1q_2\cdots q_n}{10^n-1}\left(1-10^{-mn}\right), \end{aligned}$$

dvs. att

$$\left|\frac{q_1q_2\cdots q_n}{10^n-1} - p_m\right| = \left|\frac{q_1q_2\cdots q_n}{10^n-1}10^{-mn}\right| \leq 10^{-mn}.$$

Men det innebär att för fixt $n > 0$ kan skillnaden mellan $\frac{q_1q_2\cdots q_n}{10^n-1}$ och p_m göras godtyckligt liten genom att välja större m. Alltså motsvarar decimalexpansionen det rationella talet

$$p = \frac{q_1q_2\cdots q_n}{10^n-1}.$$

Därmed har det visats att alla periodiska decimalexpanisoner är lika med ett rationellt tal. □

1.2 Kapitel 12

Definiera begreppet "Lipschitzkontinuitet".

En funktion f är *Lipschitzkontinuerlig* på ett intervall I om

$$|f(x) - f(y)| \leq L|x-y|, \quad \forall x, y \in I, \tag{3}$$

där $L > 0$ kallas *Lipschitzkonstanten.* □

Visa att summan av två Lipschitzkontinuerliga funktioner är Lipschitzkontinuerlig.

Vi har att funktionerna $f_1(x)$ och $f_2(x)$ är Lipschitzkontinuerliga med Lipschitzkonstanter L_1 resp. L_2. Utifrån (3) bildar vi

$$\begin{aligned} |(f_1(x)+f_2(x)) - (f_1(y)+f_2(y))| &= |(f_1(x)-f_1(y)) + (f_2(x)-f_2(y))| \\ &\leq |f_1(x)-f_1(y)| + |f_2(x)-f_2(y)| \\ &\leq L_1|x-y| + L_2|x-y| = (L_1+L_2)|x-y|, \end{aligned}$$

genom användning av TO (samt att f_1 och f_2 var för sig är Lipschitzkontinuerliga). Alltså är summafunktionen $f_1(x) + f_2(x)$ Lipschitzkontinuerlig med $L = L_1 + L_2$.

Minns följande:

- linjärkombinationen $c_1f_1 + \cdots + c_nf_n$, för Lipschitzkontinuerliga funktioner $f_1, \ldots, f_n$ med $L_1, \ldots, L_n$, är Lipschitzkontinuerlig med $L = |c_1|L_1 + \cdots + |c_n|L_n$;
- produkten f_1f_2 av två Lipschitzkontinuerliga funktioner med L_1 resp. L_2 är Lipschitzkontinuerlig med (åtminstone) $L = M(L_1 + L_2)$ för $M = \max\{|f_1(x)|, |f_2(x)|\}$;
- kvoten f_1/f_2 mellan två Lipschitzkontinuerliga funktioner är Lipschitzkontinuerlig om $|f_2(x)| \geq m$ för något $m > 0$;
- den sammansatta funktionen $f_1 \circ f_2$ av två Lipschitzkontinuerliga funktioner, med L_1 resp. L_2 på intervallen I_1 resp. I_2, är Lipschitzkontinuerlig på I_1 med $L = L_1L_2$.

□

Vad menas med "begränsad" funktion?

En funktion f är *begränsad* på ett intervall I om det existerar en konstant M så att

$$|f(x)| \leq M, \quad \forall x \in I. \tag{4}$$

Lipschitzkontinuerliga funktioner är alltid begränsade: Låt $I = [a, b]$ med $|I| = |b - a|$. Det följer för fixt y att

$$|f(x) - f(y)| \leq L|x-y| \implies |f(x)| \leq |f(y)| + L|x-y| \leq |f(y)| + L|I|,$$

ty $|x - y| \leq |I|$. Alltså är $|f(x)|$ begränsad för alla x (även om vi inte känner $|f(y)|$ vet vi att konstanten är begränsad). □

1.3 Kapitel 13

Definiera begreppen "gränsvärde" och "konvergens".

Gränsvärdet för en följd $\{a_n\}_{n=1}^{\infty}$ är lika med A, vilket betecknas

$$\lim_{n\to\infty} a_n = A, \tag{5}$$

om det för varje $\epsilon > 0$ finns ett $N \in \mathbb{N}$ så att

$$|a_n - A| \leq \epsilon, \quad \forall n \geq N. \tag{6}$$

En *konvergent* följd är en följd som har gränsvärde. Motsatsen kallas *divergent* följd. □

Visa att om $\{a_n\}$ är konvergent och f Lipschitzkontinuerlig så måste $\{f(a_n)\}$ vara konvergent med $\lim_{n\to\infty} f(a_n) = f\left(\lim_{n\to\infty} a_n\right)$.

En given följd är konvergent och vi undrar huruvida den Lipschitzkontinuerliga funktionen är detsamma med talen från följden som input? Svaret är "Ja!" eftersom

$$\left| f(a_m) - f\left(\lim_{n\to\infty} a_n\right) \right| \leq L \left| a_m - \lim_{n\to\infty} a_n \right|,$$

där HL kan göras godtyckligt litet genom val av större m. Alltså följer från (5)–(6) att gränsvärdet för $\{f(a_m)\}$ existerar som

$$\lim_{n\to\infty} f(a_n) = f\left(\lim_{n\to\infty} a_n\right).$$

□

1.4 Kapitel 14

Visa att kvadratroten ur 2 inte är ett rationellt tal.

Vi vet att ett godtyckligt heltal N kan primtalfaktoriseras

$$N = p_1^a \cdot p_2^b \cdot p_3^c \cdots,$$

för p_i primtal och $a, b, c, \ldots$ heltalsexponenter. Genomför ett motsägelsebevis: Antar först att $\sqrt{2}$ är ett rationellt tal, dvs. att det kan skrivas

$$\sqrt{2} = \frac{r}{s},$$

för heltal r, s sådana att alla gemensamma faktorer i kvoten förkortats bort. Men då är

$$\sqrt{2}r = s \implies 2r^2 = s^2,$$

vilket innebär att HL måste innehålla åtminstone en faktor 2 (annars kan inte likheten gälla). Det finns faktiskt minst två faktorer 2 (s innehöll från början endast heltal), dvs. att vi kan skriva $s = 2\bar{s}$. Följer att

$$2r^2 = 4\bar{s}^2 \implies r^2 = 2\bar{s}^2,$$

varför även r — via samma resonemang som förut — innehåller minst en faktor 2. Men det strider mot antagandet (r, s har gemensamma faktorer trots allt)! Har därmed visat att $\sqrt{2}$ inte är ett rationellt tal. □

Redogör för hur bisektionsalgoritmen fungerar.

Bisektionsalgoritmen används för att lösa ekvationer på formen $f(x) = 0$ genom instängning av nollställen. Givet ett startintervall, som säkert har ett nollställe (se Kapitel 16 och Bolzanos sats) evalueras funktionen i dess mittpunkt. Finner på så sätt vilken hälft av intervallet som har nollstället. Förkastar den andra och bildar ett nytt sökintervall. Upprepar proceduren tills dess att instängningen hamnat tillräckligt nära roten.

Bisektionsschema

1. välj startintervall $I_0 = [x_0, X_0]$ så att $f(x_0)f(X_0) < 0$. Låt $i = 1$;
2. evaluera f i mittpunkten $\overline{x}_i = (x_{i-1} + X_{i-1})/2$. Om
   - $f(\overline{x}_i) <$ TOL avbryt,
   - $f(\overline{x}_i) < 0$ sätt $x_i = \overline{x}_i$ och $X_i = X_{i-1}$,
   - $f(\overline{x}_i) > 0$ sätt $x_i = x_{i-1}$ och $X_i = \overline{x}_i$;
3. uppdatera $i = i + 1$ och fortsätt från 2.

Vi säger "bisektion" därför att instängningen sker genom successiv halvering av intervallet. □

Definiera begreppet "Cauchyföljd".

En följd $\{x_i\}_{i=1}^{\infty}$ är en *Cauchyföljd* om det för varje $\epsilon > 0$ finns ett $N \in \mathbb{N}$ så att

$$|x_i - x_j| \leq \epsilon, \quad \forall i, j \geq N. \tag{7}$$

Med andra ord ska talen $x_1, x_2, \ldots$ i följden bli mer och mer lika varandra "ju längre bort vi tittar".

Exempel. För följden $\left\{\frac{1}{n^2}\right\}$, om vi antar $j \geq i$, fås

$$\left| \frac{1}{i^2} - \frac{1}{j^2} \right| = \left| \frac{j^2 - i^2}{i^2 j^2} \right| \leq \frac{j^2}{i^2 j^2} = \frac{1}{i^2},$$

som vi inser blir godtyckligt litet när $i \to \infty$. Särskilt är skillnaden mellan talen i följden mindre än ϵ när

$$\frac{1}{N^2} \leq \epsilon \implies N \geq \frac{1}{\sqrt{\epsilon}}.$$

Alltså satisfierar följden (7) för (åtminstone) $j \geq i \geq N = 1/\sqrt{\epsilon}$ och är en Cauchyföljd. □

1.5 Kapitel 15

Cauchyföljder och reella tal?

Vet att en oändlig periodisk decimalutveckling motsvarar ett rationellt tal, men alla utvecklingar är inte periodiska. Vi kallar en icke-periodisk decimalutveckling $x = \pm p_m \cdots p_0.q_1 q_2 \cdots$ för ett *reellt tal* och skriver

$$x = \lim_{i\to\infty} x_i,$$

där $\{x_i\}_{i=1}^{\infty}$ får vara den trunkerade utvecklingen av x. Följden är en Cauchyföljd eftersom

$$|x_i - x_j| \leq 10^{-i}, \quad \forall j > i, \tag{8}$$

då i betecknar antalet lika decimaler mellan det i:te talet i följden och samtliga efterföljande.

Alla Cauchyföljder kan, eventuellt efter strykning av termer (för att garantera konvergenshastigheten) och omindexering, göras om till en decimalexpansion som uppfyller (8). Talen i en Cauchyföljd blir ju successivt mer och mer lika varandra! Trots att den nya och den gamla följden är olika har de ändå samma gränsvärde. Att ett sådant existerar inses från att talen i följden, oavsett om de blir större eller mindre när nya decimaler tillkommer (beroende på om x är positivt eller negativt), alltid är begränsade. Därmed kan inte följden vara divergent (måste istället vara konvergent).

Slutsats: Cauchyföljder kan skrivas som decimalutvecklingar vars gränsvärden är reella tal.

Anmärkning. Alla decimalutvecklingar, antingen periodiska eller icke-periodiska, är reella tal. De reella talen omfattar alltså såväl rationella som irrationella tal.

Exempel. Vi har att $\sqrt{2}$ är gränsvärdet för

$$\{1.000000, 1.250000, 1.375000, 1.406250, 1.414184, 1.414199, 1.414207, \ldots\} \tag{9}$$

där successivt fler och fler decimaler i utvecklingen överensstämmer. Följden (9) är den som genereras i undre gräns när bisektionsalgoritmen tillämpas på ekvationen $f(x) = x^2 - 2 = 0$ [AMBS, sek. 14.3]. Notera att dubbla förekomster av tal strukits. Vi kan stryka ännu fler tal för att skriva

$$\{1.406250, 1.414184, 1.414199, 1.414207, \ldots\}$$

och skillnaden mellan $x_2 = 1.414184$ och senare tal i följden är aldrig större än 10^{-2}. Exempelvis är

$$|x_2 - x_4| = |1.414184 - 1.414207| = 0.000023 < 10^{-2},$$

Sedan $|x_3 - x_4| = 0.000008 \leq 10^{-3}$. Gränsvärdet är fortfarande $\sqrt{2}$. □

Visa att om $\{x_i\}$ är en Cauchyföljd och f är Lipschitzkontinuerlig så är $f(x_i)$ en Cauchyföljd.

Har att

$$|f(x_i) - f(x_j)| \leq L|x_i - x_j|,$$

men då HL kan göras godtyckligt litet, innebär det att differensen i VL uppfyller (7). Alltså är $\{f(x_i)\}$ också en Cauchyföljd. □

1.6 Kapitel 16

Sats 1.1 (BOLZANO). *Om $f : [a, b] \to \mathbb{R}$ är Lipschitzkontinuerlig och $f(a)f(b) < 0$ existerar ett reellt tal $\overline{x} \in [a, b]$ så att $f(\overline{x}) = 0$.*

(Alternativt: Med f som ovan hittas säkert ett nollställe via bisektion.)

Bevis. Söker $\overline{x}$ med bisektion. Efter i iterationer har vi

$$0 \leq x_i - X_i \leq 2^{-i}|x_0 - X_0|,$$

p.g.a. algoritmens konstruktion (halverar sökintervallet i varje steg). Vet att $x_i \leq x_j \leq X_j \leq X_i$ för $j \geq i$. Därför måste även

$$|x_i - x_j| \leq 2^{-i}|x_0 - X_0|,$$

vilket ger $\{x_i\}$ som en Cauchyföljd. Alltså konvergerar bisektionsalgoritmen mot ett reellt tal $\overline{x}$ så att

$$|x_i - \overline{x}| \leq 2^{-i}|x_0 - X_0|, \quad |X_i - \overline{x}| \leq 2^{-i}|x_0 - X_0|.$$

Måste också visa att gränsvärdet är ett nollställe, dvs. att $\overline{x}$ löser ekvationen $f(x) = 0$. Känner till

$$f(\overline{x}) = f\left(\lim_{i\to\infty} x_i\right) = \lim_{i\to\infty} f(x_i),$$

eftersom f är Lipschitzkontinuerlig, så om $\lim_{i\to\infty} f(x_i) = 0$ är vi i hamn. Men om

$$|f(x_i) - f(X_i)| \leq L|x_i - X_i| \leq L2^{-i}|x_0 - X_0|,$$

och $f(x_i)f(X_i) < 0$ (faktorerna har olika tecken), vilket medför att $|f(x_i)| \leq |f(x_i) - f(X_i)|$, fås

$$|f(x_i)| \leq L2^{-i}|x_0 - X_0|.$$

Då kan ju HL göras godtyckligt litet, varför $\lim_{i\to\infty} f(x_i) = 0$, och vi är klara. □

Sats 1.2 (MELLANLIGGANDE VÄRDEN). *Om $f : [a, b] \to \mathbb{R}$ är Lipschitzkontinuerlig antar f alla värden mellan $f(a)$ och $f(b)$ (åtminstone en gång).*

Bevis. Tillämpa Bolzanos sats på funktionen $g(x) = f(x) - y = 0$, där y antar alla värden mellan $f(a)$ och $f(b)$ (genom y väljs alla mellanliggande värden till nollställen för g, som enligt Bolzanos sats kan hittas via bisektion). □

1.7 Kapitel 19

Sats 1.3 (KONTRAKTIONSAVBILDNING). *Om $g : \mathbb{R} \to \mathbb{R}$ är Lipschitzkontinuerlig med $L < 1$ (kontraktion), har g en unik fixpunkt $\overline{x} \in \mathbb{R}$, och varje följd $\{x_i\}_{i=1}^{\infty}$ som genereras via fixpunktsiteration konvergerar mot $\overline{x}$.*

Bevis. Beviset utförs i tre steg:

1. visar att $\{x_i\}_{i=1}^{\infty}$ från fixpunktsiteration är en Cauchyföljd (så att schemat konvergerar och ett gränsvärde $\overline{x}$ existerar);
2. visar att $\overline{x} = \lim_{i\to\infty} x_i$ är en fixpunkt;

3. visar att $\overline{x}$ är unik.

Fixpunktsiteration genererar Cauchyföljd

Nya gissningar vid fixpunktsiteration bildas som $x_{k+1} = g(x_k)$. Därför blir

$$x_{k+1} - x_k = g(x_k) - g(x_{k-1}) \leq L|x_k - x_{k-1}|, \tag{10}$$

när g är Lipschitzkontinuerlig. På samma sätt måste

$$x_k - x_{k-1} \leq L|x_{k-1} - x_{k-2}|, \tag{11}$$

och med (11) insatt i (10)

$$x_{k+1} - x_k \leq L^2|x_{k-1} - x_{k-2}|.$$

Upprepas argumentet fås till slut

$$x_{k+1} - x_k \leq L^k|x_1 - x_0|, \tag{12}$$

vilket visar sig vara ett användbart resultat. Men först konstateras att för $j > i$ gäller

$$|x_i - x_j| = |x_i - x_{i+1} + x_{i+1} - x_{i+2} + x_{i+2} + \cdots + x_{j-1} - x_j|,$$

eller, via TO, att

$$|x_i - x_j| \leq |x_i - x_{i+1}| + |x_{i+1} - x_{i+2}| + \cdots + |x_{j-1} - x_j| = \sum_{k=i}^{j-1} |x_k - x_{k+1}|.$$

Nu, genom (12), fås

$$|x_i - x_j| \leq \sum_{k=i}^{j-1} L^k|x_1 - x_0| = |x_1 - x_0| \sum_{k=i}^{j-1} L^k,$$

där HL är en geometrisk summa. Vet då

$$\sum_{k=i}^{j-1} L^k = L^i \left(1 + L + L^2 + \cdots + L^{j-i-1}\right) = L^i \frac{1 - L^{j-i}}{1 - L},$$

och med $L < 1$ måste $0 \leq 1 - L^{j-i} \leq 1$, eller $L^i(1 - L^{j-i}) \leq L_i$, så att

$$|x_i - x_j| \leq \frac{L^i}{1 - L}|x_1 - x_0|.$$

Men eftersom L^i kan göras godtyckligt liten när i väljs större följer att $\{x_i\}_{i=1}^{\infty}$ är en Cauchyföljd med ett gränsvärde. Vi kallar det $\overline{x}$.

Gränsvärdet är en fixpunkt

När g är Lipschitzkontinuerlig ($\star$) följer

$$g(\overline{x}) = g\left(\lim_{i\to\infty} x_i\right) \overset{\star}{=} \lim_{i\to\infty} g(x_i) = \lim_{i\to\infty} x_{i+1} = \overline{x},$$

ty för fixpunktsiteration blir $x_{i+1} = g(x_i)$ och givetvis är

$$\lim_{i\to\infty} x_{i+1} = \lim_{i\to\infty} x_i.$$

Med andra ord måste $\overline{x}$ vara en fixpunkt för g.

Entydighet

Antag att det finns två fixpunkter x och y så att $g(x) = x$ samt $g(y) = y$. Men eftersom g är en kontraktion måste

$$|x - y| = |g(x) - g(y)| \leq L|x - y|, \tag{13}$$

så att enda möjligheten är $x = y$ (annars kan (13) inte vara sant för $L < 1$). Alltså blir fixpunkten unik.

Äntligen är vi klara! □

1.8 Kapitel 22

Sats 1.4 (FAKTORSATSEN)**.** *Om $\overline{x}$ är ett nollställe till polynomekvationen $f(x) = 0$ så är polynomet $x - \overline{x}$ en faktor i $f(x)$, dvs.*

$$f(x) = (x - \overline{x})g(x),$$

för något polynom $g(x)$.

Sats 1.5 (ALGEBRANS FUNDAMENTALSATS)**.** *Om grad $p = n$ har (den möjligt komplexa) polynomekvationen $p(z) = 0$ precis n nollställen, när varje nollställe räknas så många gånger som dess multiplicitet anger.*

1.9 Kapitel 23

Definiera begreppet "derivata".

Funktionen $f(x)$ är *deriverbar* på ett intervall I, om f är deriverbar för alla punkter $\overline{x} \in I$, så att för x nära $\overline{x}$

$$f(x) = f(\overline{x}) + m(\overline{x})(x - \overline{x}) + E_f(x, \overline{x}), \tag{14}$$

där $|E_f(x, \overline{x})| \leq K_f(\overline{x})(x - \overline{x})^2$. I så fall är derivatan $f'(\overline{x}) = m(\overline{x})$. Vi känner igen tangenten

$$\tilde{f}_{\overline{x}}(x) = f(\overline{x}) + f'(\overline{x})(x - \overline{x}),$$

som linjäriseringen av f i punkten $\overline{x}$. Alltså säger (14) att felet $E_f(x, \overline{x}) = f(x) - \tilde{f}_{\overline{x}}(x)$ — avvikelsen mellan funktionsgrafen och tangenten — ska vara kvadratiskt litet i $x - \overline{x}$. Särskilt är den linjära approximationen god för x nära $\overline{x}$ (däremot behöver den inte vara det på längre avstånd). □

Sats 1.6. *Om $f(x)$ är likformigt differentierbar på intervallet $I = [a,b]$ och det finns en konstant L så att*

$$|f'(x)| \leq L, \quad \forall x \in I,$$

är f Lipschitzkontinuerlig med Lipschitzkonstant L.

Bevis. Vet från definitionen av likformigt differentierbar funktion att

$$f(x) = f(y) + f'(y)(x-y) + E_f(x,y),$$

där $|E_f(x,y)| \leq K_f(x-y)^2$ (för någon konstant K_f). Kan då skriva

$$f(x) = f(y) + (L + K_f|x-y|)|x-y|,$$

vilket visserligen ger att f är Lipschitzkontinuerlig, fast med den större Lipschitzkonstanten $\overline{L} = L + K_f|b-a|$. Men vad händer om vi istället betraktar f på ett mindre delintervall I_δ av I med längden δ? Här gäller ju

$$f(x) = f(y) + (L + K_f|x-y|)|x-y| \leq f(y) + (L + K_f\delta)|x-y|, \tag{15}$$

ty $|x-y| \leq \delta$ på I_δ. Om $\delta \to 0$ fås att $L + K_f\delta \to L$. Alltså är Lipschitzkonstanten L på I_δ. Låt nu I delas in i N delintervall genom $x = x_0 < x_1 < \cdots < x_{N-1} < X_N = y$, vars längder $x_i - x_{i-1} \leq \delta$. Får att

$$|f(x) - f(y)| = \left|\sum_{i=1}^{N}(f(x_i) - f(x_{i-1}))\right| \leq \sum_{i=1}^{N}|f(x_i) - f(x_{i-1})|,$$

via TO. Utnyttja sedan (15) på varje delintervall

$$\sum_{i=1}^{N}|f(x_i) - f(x_{i-1})| \leq (L + K_f\delta)\sum_{i=1}^{N}|x_i - x_{i-1}| = (L + K_f\delta)|x-y|,$$

ty

$$\sum_{i=1}^{N}|x_i - x_{i-1}| = |x-y|$$

(summan av delintervallens längder motsvarar hela intervallets längd). Men eftersom indelningen av I kan göras godtyckligt tät (när $\delta \to 0$) måste

$$|f(x) - f(y)| \leq L|x-y|,$$

och det följer att f är Lipschitzkontinuerlig överallt med Lipschitzkonstant L. □

1.10 Kapitel 25

Newtons metod.

Antar att $g(x)$ är en Lipschitzkontinuerlig funktion på $I = [a,b]$ med $L < 1$ (kontraktion). Då har g en unik fixpunkt $\overline{x} \in I$ så att $g(\overline{x}) = \overline{x}$ (kontraktionsavbildningssatsen). Kan bestämmas via fixpunktsiteration: $x_{i+1} = g(x_i)$, $i = 1, 2, \ldots$ (givet en startgissning x_1).

Hur snabbt konvergerar fixpunktsiteration?

För x nära $\overline{x}$ har vi följande kvadratiska approximation av g:

$$g(x) = g(\overline{x}) + g'(\overline{x})(x-\overline{x}) + \tfrac{1}{2}g''(\overline{x})(x-\overline{x})^2 + E_f(x,\overline{x}),$$

där $|E_f(x,\overline{x})| \leq K_g(\overline{x})|x-\overline{x}|^3$. Låt nu $x = x_i$ samt $e_i = x_i - \overline{x}$ (rotfelet för fixpunktsiteratet x_i) och vi får

$$e_{i+1} = x_{i+i} - \overline{x} = g(x_i) - g(\overline{x}) = g'(\overline{x})(x_i - \overline{x}) + \tfrac{1}{2}g''(\overline{x})(x_i-\overline{x})^2 + E_g(x_i,\overline{x}).$$

Men för $g'(\overline{x}) \neq 0$ kan vi uppskatta

$$|e_{i+1}| \approx |g'(\overline{x})||e_i|,$$

eftersom den linjära termen blir dominant. Vi säger därför att fixpunktsiteration har *linjär konvergens*[1] (upp till en faktor $|g'(\overline{x})|$).

Kan vi få bättre?

Ja — om $g'(\overline{x}) = 0$ blir istället

$$|e_{i+1}| \approx |g''(\overline{x})||e_i|^2,$$

ty den kvadratiska termen är dominant. Det innebär *kvadratisk konvergens.*

Frågeställning:

Vill lösa $f(x) = 0$ som ett fixpunktsproblem via omskrivningen

$$g(x) = x - \alpha(x)f(x) = x.$$

Idé:

Välj $\alpha(\overline{x})$ så att $g'(\overline{x}) = 0$ för att få kvadratisk konvergens. Vi prövar

$$g'(\overline{x}) = 1 - \alpha'(\overline{x})f(\overline{x}) - \alpha(\overline{x})f'(\overline{x}) = 0,$$

och eftersom $f(\overline{x}) = 0$ fås

$$\alpha(\overline{x}) = \frac{1}{f'(\overline{x})},$$

för $f'(\overline{x}) \neq 0$.

Slutsats:

[1] Konvergenshastigheten är egentligen lokal då vi antagit x nära $\overline{x}$.

Med $\alpha(x) = \frac{1}{f'(x)}$ ges iterationsschemat

$$x_{i+1} = g(x_i) = x_i - \frac{f(x_i)}{f'(x_i)}, \quad i = 1, 2, \ldots, \tag{16}$$

givet en startgissning x_1. Newtons metod är alltså ett specialfall av fixpunktsiteration (väljer ett smart $\alpha(x)$) som ger förbättrade konvergensegenskaper.

BEVIS (kvadratisk konvergens)

Givet $f'(x_i) \neq 0$ har vi från (16)

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

och vet från linjäriseringen av $f(\overline{x}) = 0$ kring x_i

$$\begin{aligned} f(\overline{x}) &= f(x_i) + f'(x_i)(\overline{x} - x_i) + E_f(\overline{x}, x_i) = 0, \\ f(x_i) &= -f'(x_i)(\overline{x} - x_i) - E_f(\overline{x}, x_i), \end{aligned}$$

så att

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

med

$$|x_{i+1} - \overline{x}| = \left| \frac{E_f(\overline{x}, x_i)}{f'(x_i)} \right| \leq \frac{K_f(x_i)}{|f'(x_i)|}(x_i - \overline{x})^2,$$

ty $|E_f(\overline{x}, x_i)| \leq K_f(x_i)(x_i - \overline{x})^2$ enligt derivatans definition. Med en startgissning x_1 tillräckligt nära $\overline{x}$ fås m.a.o. kvadratisk konvergens. Ett bra x_1 kan t.ex. hittas via bisektionsalgoritmen.

Geometrisk tolkning

Vi bildar tangenten $\tilde{f}_{x_i}(x)$ till $f(x)$ i x_i som

$$\tilde{f}_{x_i}(x) = f(x_i) + f'(x_i)(x - x_i),$$

och låter x_{i+1} vara x-värdet för tangenten skärning med x-axeln

$$\tilde{f}_{x_i}(x_{i+1}) = f(x_i) + f'(x_i)(x_{i+1} - x_i) = 0,$$

varpå det följer

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

Men detta är ju Newtons metod! Alltså motsvaras (16) geometriskt av att lösa det linjäriserade problemet $\tilde{f}_{x_i}(x) = 0$ iterativt (istället för det ursprungliga $f(x) = 0$).

Hur stort är felet?

Kan vi säga något om rotfelet $x_i - \overline{x}$ givet residualen $f(x_i)$? (Minns att $f(x_i) = 0$ för $x_i = \overline{x}$.) Ja, det kan man, ty

$$f(x_i) = f(x_i) - f(\overline{x}) = f'(\overline{x})(x_i - \overline{x}) + E_f(x_i, \overline{x}),$$

och med $f'(\overline{x}) \neq 0$ fås

$$x_i - \overline{x} \approx \frac{f(x_i)}{f'(\overline{x})}, \tag{17}$$

upp till en restterm $E_f(x_i, \overline{x})/f'(\overline{x})$ (som vi vet är kvadratisk liten i $x_i - \overline{x}$). Alltså är rotfelet proportionellt mot residualen för x_i nära $\overline{x}$.

Anmärkning. Märk att om $f'(\overline{x}) \approx 0$ så kan rotfelet vara stort trots att residualen är liten (vi kan bli lurade). Problemet sägs i så fall vara illa konditionerat (ill-conditioned).

Används (17) i (16) ges

$$x_{i+1} \approx x_i - (x_i - \overline{x}) \implies |x_{i+1} - x_i| \approx |x_i - \overline{x}|, \tag{18}$$

dvs. att rotfelet i $x_i - \overline{x}$ kan uppskattas m.h.a. efterföljande iterat som $x_{i+1} - x_i$.

När avbryts iterationsschemat?

Då den approximativa lösningen är tillräckligt bra! Två naturliga stoppkriterier är

$$\left| \frac{f(x_i)}{f'(x_i)} \right| \leq \text{TOL}, \quad |x_{i+1} - x_i| \leq \text{TOL},$$

från (17) resp. (18). Alternativt kan residualen utnyttjas. TOL är en given tolerans. □