



Forelesning II

IMAT2150: Høsten 2023

Kent Holing, NTNU (23. august 2023)

VIKTIG RETTELSE ang. øvinger:

Det er nok at **70%** er godkjent for å kunne gå opp til eksamen (og ikke 80% som ble opplyst på forrige forelesning).

VIKTIG RETTELSE ang. øvinger:

Det er nok at **70%** er godkjent for å kunne gå opp til eksamen (og ikke 80% som ble opplyst på forrige forelesning).

Kort tilbakeblikk fra forrige forelesning

Refleksjon:

Noen læringspunkter basert på erfaring fra første forelesning.

LÆRINGSMÅL (forrige forelesning)

- Evaluering av polynom (bla. Horners metode) (S: 0.1)
- Binære tall - konvertering mellom desimaltall og binærtall (begge veier) (S: 0.2)
- Andre tallsystem (S: 0.2)
- Flyttall: Representasjon, formater og beregninger (S: 0.3)

Merk at for $0.7 = 7/10 = 7/(2 \times 5)$ er **nevneren ikke en toer-potens**.

Fractional part. Convert $(0.7)_{10}$ to binary by reversing the preceding steps. Multiply by 2 successively and record the integer parts, moving away from the decimal point to the right.

$$\begin{aligned} .7 \times 2 &= .4 + 1 \\ .4 \times 2 &= .8 + 0 \\ .8 \times 2 &= .6 + 1 \\ .6 \times 2 &= .2 + 1 \\ .2 \times 2 &= .4 + 0 \\ .4 \times 2 &= .8 + 0 \\ &\vdots \end{aligned}$$

Notice that the process repeats after four steps and will repeat indefinitely exactly the same way. Therefore,

$$(0.7)_{10} = (.1011001100110\dots)_2 = (.1\overline{0110})_2.$$

Merk at for $0.7 = 7/10 = 7/(2 \times 5)$ er **nevneren ikke en toer-potens**.

Fractional part. Convert $(0.7)_{10}$ to binary by reversing the preceding steps. Multiply by 2 successively and record the integer parts, moving away from the decimal point to the right.

$$\begin{aligned} .7 \times 2 &= .4 + 1 \\ .4 \times 2 &= .8 + 0 \\ .8 \times 2 &= .6 + 1 \\ .6 \times 2 &= .2 + 1 \\ .2 \times 2 &= .4 + 0 \\ .4 \times 2 &= .8 + 0 \\ &\vdots \end{aligned}$$

Notice that the process repeats after four steps and will repeat indefinitely exactly the same way. Therefore,

$$(0.7)_{10} = (.1011001100110\dots)_2 = (.1\overline{0110})_2.$$

La oss kontrollere svaret!

Det nedenfor er en rettelse av det som ble vist på forrige forelesning!

La $x = (0.1\overline{0110})_2$.

Da er $2x = 1.\overline{0110} = 1 + \overline{0110}$.

Med $y = \overline{0110}$ er $2^4 y = 0110 + y$, så $y = 6/15 = 2/5$.

M.a.o. er $2x = 1 + y = 1 + 2/5 = 7/5$, og $x = 7/10 = 0.7$, **som skulle vises**.

LÆRINGSMÅL

- Tap av signifikante sifre (S: 0.4)
- Halveringsmetoden (S: 1.1)
- Fikspunktiterasjon (FPI) (S: 1.2)

TAP AV SIGNIFIKANTE SIFRE

Tap av signifikans

$$\begin{aligned} a &= 252.164 \quad \Rightarrow 6 \text{ signifikante siffer} \\ b &= 252.163 \end{aligned}$$

$$a+b = 504.327 \quad - 6 \text{ sign. siffer}$$

$$a-b = 0.001 \quad - 1 \text{ sign. siffer}$$

```
$ python3
Python 3.6.9 (default, Mar 10 2023, 16:46:00)
[GCC 8.4.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> 252.164-252.163
0.0009999999999763531
>>> 252.164+252.163
504.327
>>>
```

Ek. $\sqrt{4 + 2^{-50}} - 2$

Python: $\sqrt{4 + 2 \times 10^{-50}} - 2$
0

$$\left(\sqrt{4 + 2^{-50}} - 2 \right) \cdot \frac{\left(\sqrt{4 + 2^{-50}} + 2 \right)}{\left(\sqrt{4 + 2^{-50}} + 2 \right)} = \frac{4 + 2^{-50} - 4}{\sqrt{4 + 2^{-50}} + 2}$$

$$\approx 2.2204 \cdot 10^{-16}$$

Ek. $\sqrt{4 + 2^{-50}} - 2$

Python: $\text{sqrt}(4 + 2^{**}(-50)) - 2$
0

$$\frac{(\sqrt{4 + 2^{-50}} - 2) \cdot (\sqrt{4 + 2^{-50}} + 2)}{(\sqrt{4 + 2^{-50}} + 2)} = \frac{4 + 2^{-50} - 4}{\sqrt{4 + 2^{-50}} + 2}$$

$$\approx 2.2204 \cdot 10^{-16}$$

```
import math
expr1=math.sqrt(4+2**(-50))-2
print(expr1)
expr2=(4+2**(-50)-4)/(math.sqrt(4+2**(-50))+2)
print(expr2)
```

```
0.0
2.22044604925e-16
```

Eksempel 0.6 (s.17–18) $x^2 + 9^{12} x - 3 = 0$ i læreboka.

Brukes **standard formel** for løsningen av andregradsligningen så gir MATLAB med **minus foran rottegnet** -2.824×10^{11} , som er korrekt med 4 signifikante sifre.

Med **pluss foran rottegnet** gir MATLAB (og Python) 0, som opplagt er totalt feil!

Eksempel 0.6 (s.17–18) $x^2 + 9^{12} x - 3 = 0$ i læreboka.

Brukes **standard formel** for løsningen av andregradsligningen så gir MATLAB med **minus foran rottegnet** -2.824×10^{11} , som er korrekt med 4 signifikante sifre.

Med **pluss foran rottegnet** gir MATLAB (og Python) 0, som opplagt er totalt feil!

En modifisert formel (neste lysark) gir 1.1062×10^{-11} som også er korrekt med 4 signifikante sifre.

IMAT2150: Tap av signifikante sifre

Stabil løsning av andregradsligningen

Eksempel 0.6 (s.17–18) $x^2 + 9^{12}x - 3 = 0$ i læreboka.

Brukes **standard formel** for løsningen av andregradsligningen så gir MATLAB med **minus foran rottegnet** -2.824×10^{11} , som er korrekt med 4 signifikante sifre.

Med **pluss foran rottegnet** gir MATLAB (og Python) 0, som opplagt er totalt feil!

En modifisert formel (neste lysark) gir 1.1062×10^{-11} , som også er korrekt med 4 signifikante sifre.

Programsystemet Mathematica® (med utvidet presisjon) gir:

```
Mathematica 13.0.1 Kernel for Linux x86 (64-bit)
Copyright 1988-2022 Wolfram Research, Inc.

In[1]:= Solve[x^2 + 9^12 x - 3 == 0, x]

Out[1]= {{x -> 
$$\frac{-282429536481 - \text{Sqrt}[79766443076872509863373]}{2}$$
}, {x -> 
$$\frac{6}{282429536481 + \text{Sqrt}[79766443076872509863373]}$$
}}

In[2]:= N[%, 20]

Out[2]= {{x -> 
$$-2.82429536481000000000 \cdot 10^{11}$$
}, {x -> 
$$1.0622118484416449309 \cdot 10^{-11}$$
}}
```

Løse 2. gradslikninger uten tap av signifikans.

Restrukturert abc-formel.

For $ax^2 + bx + c = 0$ med $b > 0$:

$$x_1 = -\frac{(b + \sqrt{b^2 - 4ac})}{2a}$$

$$x_2 = -\frac{2c}{b + \sqrt{b^2 - 4ac}}$$

For $ax^2 + bx + c = 0$ med $b < 0$:

$$x_1 = -\frac{b - \sqrt{b^2 - 4ac}}{2a}$$

```
quadsolver[a_,b_,c_]:=Module[{delta,q,x1,x2},
delta=b^2-4a c;
q=- (b+Sign[b] Sqrt[delta])/2;
x1=q/a;
x2=c/q;
Return[{x1,x2}]];
{a,b,c}={1,9^12,-3}; (** from SAUER Examø 0.6 **)
roots=quadsolver[a,b,c];
Print[roots//N];

{-2.82429536481*^11, 1.062211848441645*^-11}
```

Dette avsnittet i SAUER vil ikke bli gått igjennom fortløpende i sin helhet, men jeg vil repetere stoff herfra første gang det er aktuelt å bruke det i forelesningene.¹

¹Dette skal jo såvidt meg bekjent være kjent stoff fra tidligere.

SAUER KAPITTEL 1

Løsning av ligninger med en ukjent

Følgende setning er intuitivt riktig.^{1,2}

Skjæringssetningen ("Intermediate Value Theorem")

Dersom $f(x)$ er kontinuerlig på intervallet $[a, b]$ og $f(a) \times f(b) < 0$, så har ligningen $f(x) = 0$ minst én rot i (a, b) .

¹Hva skjer når du tegner grafen til en kontinuerlig (sammenhengende) funksjon fra et punkt under x-aksen til et punkt over x-aksen uten å løfte blyanten?

²Se Theorem 0.4 SAUER 0.5. Setningen kan vises rigorøst matematisk, men beviset er langt fra enkelt og krever avansert kjennskap til egenskapene til de reelle tallene.

Følgende setning er intuitivt riktig.^{1,2}

Skjæringssetningen ("Intermediate Value Theorem")

Dersom $f(x)$ er kontinuertlig på intervallet $[a, b]$ og $f(a) \times f(b) < 0$, så har ligningen $f(x) = 0$ minst én rot i (a, b) .

Merk at ligningen **kan ha** røtter i intervallet (a, b) selv om $f(a)$ og $f(b)$ har **samme fortegn**, men trenger ikke ha det.

Når $f(x)$ oppfyller skjæringssetningen, så følger det at hvis $f(x)$ strengt voksende eller strengt avtagende i (a, b) , så har ligningen $f(x) = 0$ én og bare én rot i (a, b) .

¹Hva skjer når du tegner grafen til en kontinuertlig (sammenhengende) funksjon fra et punkt under x-aksen til et punkt over x-aksen uten å løfte blyanten?

²Se Theorem 0.4 SAUER 0.5. Setningen kan vises rigorøst matematisk, men beviset er langt fra enkelt og krever avansert kjennskap til egenskapene til de reelle tallene.

Skjæringssetningen danner grunnlaget for den såkalte **halveringsmetoden**.

Halveringsmetoden

Metode for å finne en rot til en funksjon.

Halveringsmetoden ligner litt på binær søk

Vi finner stadig et mindre intervall der vi har løsning.

1. Start: Et intervall $[a, b]$ som inneholder en rot.
2. Finn midtpunktet c til intervallet.
3. Bruk teoremet til å avgjøre hvilken side av c roten ligger.
4. Hopp til punkt 2 til ønsket nøyaktighet er oppnådd.
5. Svaret er et intervall.

Vi har et intervall (a, b) med midtpunkt m som indikert i den øverste boksen.
I de andre boksene er fortegnene til $f(x)$ i disse punktene merket som P (positiv) eller N (negativ) ettersom $f(x) > 0$ eller $f(x) < 0$.¹

Dette gir oss informasjon hvor **neste søkeintervall** for halveringsmetoden er.

(Dvs. behold intervallet der funksjonen skifter fortegn.)

$ a m b $

$ P P N $

r er i (m, b) : sett $a = m$, behold b

$ N P P $

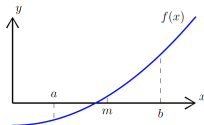
r er i (a, m) : sett $b = m$, behold a

$ P N N $

r er i (a, m) : sett $b = m$, behold a

$ N N P $

r er i (m, b) : sett $a = m$, behold b



¹Er $f(m) = 0$, så stopper halveringsmetoden da roten m er lokalisert.

Bestem kvadratoten av 3.¹

Iterations	a	b	t	f(a)	f(b)	f(t)
1	1	2	1.5	-2	1	-0.75
2	1.5	2	1.75	-0.75	1	0.062
3	1.5	1.75	1.625	-0.75	0.0625	-0.359
4	1.625	1.75	1.6875	-0.3594	0.0625	-0.1523
5	1.6875	1.75	1.7188	-0.1523	0.0625	-0.0457
6	1.7188	1.75	1.7344	-0.0457	0.0625	0.0081
7	1.7188	1.7344	1.7266	-0.0457	0.0081	-0.0189

So, at the seventh iteration, we get the final interval $[1.7266, 1.7344]$

¹Vi skal senere se på en raskere metode til å beregne kvadratrøtter på enn dette.

Les dette eksemplet selv! (Sjekk at du forstår gangen i tabellen.)

Vis at ligningen $f(x) = x^3 + x - 1 = 0$ har én og bare én rot i intervallet $(0, 1)$, og bruk halveringsmetoden til å bestemme en tilnærming til denne roten.

i	a_i	$f(a_i)$	c_i	$f(c_i)$	b_i	$f(b_i)$
0	0.0000	—	0.5000	—	1.0000	+
1	0.5000	—	0.7500	+	1.0000	+
2	0.5000	—	0.6250	—	0.7500	+
3	0.6250	—	0.6875	+	0.7500	+
4	0.6250	—	0.6562	—	0.6875	+
5	0.6562	—	0.6719	—	0.6875	+
6	0.6719	—	0.6797	—	0.6875	+
7	0.6797	—	0.6836	+	0.6875	+
8	0.6797	—	0.6816	—	0.6836	+
9	0.6816	—	0.6826	+	0.6836	+

Setningen ovenfor viser at ligningen har én og bare én rot r i intervallet $(0, 1)$.¹

Vi ser at roten r ligger i intervallet $(0.6816, 0.6826)$, dvs. $r = 0.6821 \pm 0.0005$.

¹ $f(0) < 0$, $f(1) > 0$ og $f'(x) > 0$ i $(0, 1)$.

```
1 import numpy as np
2 def bisection(f,a,b,eps):
3     it = 0
4     nr = 25 #maximum iterations allowed
5     # if f(a)*f(b) > 0:
6     if np.sign(f(a))*np.sign(f(b)) > 0:
7         return " "
8     while (b-a)/2 > eps:
9         mid = (a+b)/2
10        if f(mid)==0:
11            break
12        if f(a)*f(mid) < 0:
13            b = mid
14        else:
15            a = mid
16        if it < nr:
17            pass
18        else:
19            return mid
20        it+=1
21    return mid
```


Setning

Feilen i tilnærmingen x_c til roten r etter n halvinger oppfyller $|x_c - r| < (b - a)/2^n$.¹

Definisjon

Vi sier at x_c som tilnærmelse til r er innenfor p korrekte desimaler når feilen er mindre enn 0.5×10^{-p} .

¹For hver iterasjon halveres jo søkeintervallet.

Setning

Feilen i tilnærmingen x_c til roten r etter n halveringer oppfyller $|x_c - r| < (b - a)/2^n$.¹

Definisjon

Vi sier at x_c som tilnærmelse til r er innenfor p korrekte desimaler når feilen er mindre enn 0.5×10^{-p} .

Observasjon

Av setningen ovenfor følger:

For hver iterasjon i halveringsmetoden når den konvergerer øker antall riktige binære sifre i tilnærmelsen til r med 1.²

¹For hver iterasjon halveres jo søkeintervallet.

²Vi skal senere se at for en konvergent Newton metode doubles antall riktige binære sifre for hver iterasjon.

FIKSPUNKTITERASJON

Definisjon

Vi sier at c er et fikspunkt for $f(x)$ hvis og bare hvis $f(c) = c$.

En funksjon kan ha et, flere eller ingen fikspunkter.

Eksempelvis har $f(x) = x^2 - 3x + 3$ fikspunktene 1 og 3.

(Vis det!) (Løs ligningen $x^2 - 4x + 3 = 0$.)

Definisjon

Vi sier at c er et fikspunkt for $f(x)$ hvis og bare hvis $f(c) = c$.

En funksjon kan ha et, flere eller ingen fikspunkter.

Eksempelvis har $f(x) = x^2 - 3x + 3$ fikspunktene 1 og 3.

(Vis det!) (Løs ligningen $x^2 - 4x + 3 = 0$.)

For ligningen $g(x) = x$ er fikspunktiterasjonen definert ved $x_{i+1} = g(x_i)$ for $i = 0, 1, 2 \dots$ for en startverdi x_0 .

Definisjon

Vi sier at c er et fikspunkt for $f(x)$ hvis og bare hvis $f(c) = c$.

En funksjon kan ha et, flere eller ingen fikspunkter.

Eksempelvis har $f(x) = x^2 - 3x + 3$ fikspunktene 1 og 3.

(Vis det!) (Løs ligningen $x^2 - 4x + 3 = 0$.)

For ligningen $g(x) = x$ er fikspunktiterasjonen definert ved $x_{i+1} = g(x_i)$ for $i = 0, 1, 2 \dots$ for en startverdi x_0 .

Merk at iterasjonen trenger ikke, som vi skal se, å konvergere, men er $g(x)$ **kontinuerlig** og **iterasjonen konvergerer mot (si) r** , så er r et fikspunkt for $g(x)$:

$$g(r) = g\left(\lim_{i \rightarrow \infty} x_i\right) = \lim_{i \rightarrow \infty} g(x_i) = \lim_{i \rightarrow \infty} x_{i+1} = r$$

Ikke overraskende oppfører FPI seg forskjellig avhengig av valget av $g(x)$ for samme ligning.

Ikke overraskende oppfører FPI seg forskjellig avhengig av valget av $g(x)$ for **samme ligning**.

Eksempel på FPI

Vi skal bestemme en rot til ligningen $f(x) = x^3 + x - 1 = 0$ ved hjelp av FPI med startverdi $x_0 = 0.5$.¹

Siden $f(0) = -1 < 0$, $f(1) = 2 > 0$ og $f'(x) > 0$ i $(0, 1)$, så har ligningen én og bare én rot $r = 0.682328 \dots$ i intervallet $(0, 1)$. Dette er den eneste reelle roten til ligningen.²

¹Dette er samme eksempel som det vi brukte for halveringsmetoden.

²De andre to røttene er kompleks-konjugerte.

Ikke overraskende oppfører FPI seg forskjellig avhengig av valget av $g(x)$ for **samme ligning**.

Eksempel på FPI

Vi skal bestemme en rot til ligningen $f(x) = x^3 + x - 1 = 0$ ved hjelp av FPI med startverdi $x_0 = 0.5$.¹

Siden $f(0) = -1 < 0$, $f(1) = 2 > 0$ og $f'(x) > 0$ i $(0, 1)$, så har ligningen én og bare én rot $r = 0.682328 \dots$ i intervallet $(0, 1)$. Dette er den eneste reelle roten til ligningen.²

Vi gjør **tre omskrivninger av ligningen $f(x) = 0$ på formen $x = g(x)$** , nemlig for

1) $g_1(x) = 1 - x^3$, 2) $g_2(x) = \sqrt[3]{1-x}$ og 3) $g_3(x) = (2x^3 + 1)/(3x^2 + 1)$.³

Vi skal gi resultatet av FPI for hver av disse tilfellene på de neste lysarkene.

¹Dette er samme eksempel som det vi brukte for halveringsmetoden.

²De andre to røttene er kompleks-konjugerte.

³Vis selv at disse er gyldige omskrivninger av ligningen. Merk at $g_3(x)$ fremkommer når vi anvender **Newtons metode på $f(x) = 0$** . (Newtons metode gjennomgås senere).

Tilfelle 1: $g_1(x) = 1 - x^3$ ($r = 0.682328 \dots$)

i	x_i
0	0.50000000
1	0.87500000
2	0.33007813
3	0.96403747
4	0.10405419
5	0.99887338
6	0.00337606
7	0.99999996
8	0.00000012
9	1.00000000
10	0.00000000
11	1.00000000
12	0.00000000

FPI konvergerer ikke:

Den alternerer etter hvert mellom 0 og 1 som ikke er fikspunkter.

Vi observerer at $|g'(r)| > 1.3966 \dots > 1$.

Tilfelle 2: $g_2(x) = \sqrt[3]{1-x}$ ($r = 0.682328 \dots$)

i	x_i
0	0.50000000
1	0.79370053
2	0.59088011
3	0.74236393
4	0.63631020
5	0.71380081
6	0.65900615
7	0.69863261
8	0.67044850
9	0.69072912
10	0.67625892
11	0.68664554
12	0.67922234

i	x_i
13	0.68454401
14	0.68073737
15	0.68346460
16	0.68151292
17	0.68291073
18	0.68191019
19	0.68262667
20	0.68211376
21	0.68248102
22	0.68221809
23	0.68240635
24	0.68227157
25	0.68236807

FPI konvergerer (forholdsvis sakte).

Vi observerer at $|g'(r)| < 0.716 \dots < 1$.

Tilfelle 3: $g_3(x) = (2x^3 + 1)/(3x^2 + 1)$ ($r = 0.682328 \dots$)

i	x_i
0	0.50000000
1	0.71428571
2	0.68317972
3	0.68232842
4	0.68232780
5	0.68232780
6	0.68232780
7	0.68232780

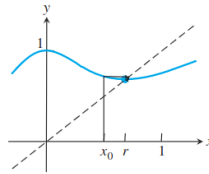
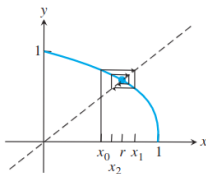
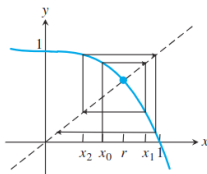
FPI konvergerer, og myere raskere enn tilfelle 2.

Vi observerer at $g'_3(r) = 0$.¹

¹Merk at $g'_3(x) = 6x f(x)/(3x^2 + 1)^2$, så $g'_3(r) = 0$ da $f(r) = 0$.

I figuren er for eksemplene våre ovenfor

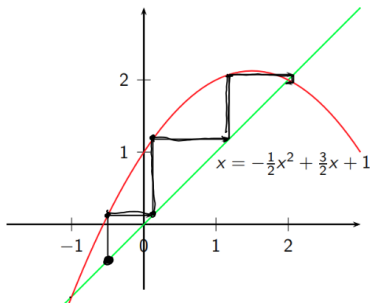
- 1) $g_1(x) = 1 - x^3$ til venstre,
- 2) $g_2(x) = \sqrt[3]{1-x}$ i midten og
- 3) $g_3(x) = (2x^3 + 1)/(3x^2 + 1)$ til høyre.



Vi sier at r er et **attraktivt fikspunkt** for $g(x)$ i FPI-en $x_{i+1} = g(x_i)$ hvis $|g'(r)| < 1$.

I figuren nedenfor er $g'(-1) = 5/2 > 1$ og $g'(2) = 1/2 < 1$, så 2 er et attraktivt fikspunkt mens -1 er det ikke.

Uansett hvordan vi velger x_0 der FPI konvergerer, så havner vi her i punktet $(2, g(2))$.



Middelverdisetningen er viktig, og vi skal bruke den flere ganger.¹

(Mean Value Theorem) Let f be a continuously differentiable function on the interval $[a, b]$. Then there exists a number c between a and b such that $f'(c) = (f(b) - f(a)) / (b - a)$. ■

Vi omskriver dette oftest som $f(b) - f(a) = f'(c) \times (b - a)$ for en c i (a, b) .

¹Dette er teorem 0.6 fra SAUER 0.5.

Teorem (FPI LOKAL konvergens)¹

La $g(x)$ være en kontinuerlig deriverbar funksjon der $g(r) = r$ og $S = |g'(r)| < 1$.

Da vil FPI konvergere **lineært** med hastighet S mot r for startverdier **tilstrekkelig nært** r .

¹Dette er teorem 1.6 i læreboka.

Teorem (FPI LOKAL konvergens)¹

La $g(x)$ være en kontinuerlig deriverbar funksjon der $g(r) = r$ og $S = |g'(r)| < 1$.

Da vil FPI konvergere **lineært** med hastighet S mot r for startverdier **tilstrekkelig nært** r .

Er $e_i = |x_i - r|$ feilen i x_i , så har vi at $e_{i+1}/e_i \rightarrow S < 1$ når $i \rightarrow \infty$.

¹Dette er teorem 1.6 i læreboka.

Teorem (FPI **LOKAL** konvergens)¹

La $g(x)$ være en kontinuerlig deriverbar funksjon der $g(r) = r$ og $S = |g'(r)| < 1$.

Da vil FPI konvergere **lineært** med hastighet S mot r for startverdier **tilstrekkelig nært** r .

Er $e_i = |x_i - r|$ feilen i x_i , så har vi at $e_{i+1}/e_i \rightarrow S < 1$ når $i \rightarrow \infty$.

Dette følger av middelverdisetningen:

Vi har jo derav $x_{i+1} - r = g(x_i) - g(r) = g'(c_i) (x_i - r)$ for c_i mellom x_i og r , som er det samme som å si at $e_{i+1} = |g'(c_i)| e_i$.

¹Dette er teorem 1.6 i læreboka.

Teorem (FPI LOKAL konvergens)¹

La $g(x)$ være en kontinuerlig deriverbar funksjon der $g(r) = r$ og $S = |g'(r)| < 1$.

Da vil FPI konvergere **lineært** med hastighet S mot r for startverdier **tilstrekkelig nært** r .

Er $e_i = |x_i - r|$ feilen i x_i , så har vi at $e_{i+1}/e_i \rightarrow S < 1$ når $i \rightarrow \infty$.

Dette følger av middelverdisetningen:

Vi har jo derav $x_{i+1} - r = g(x_i) - g(r) = g'(c_i) (x_i - r)$ for c_i mellom x_i og r , som er det samme som å si at $e_{i+1} = |g'(c_i)| e_i$.

I grensetilfellet sier dette at $e_{i+1}/e_i \rightarrow |g'(r)| = S$ når $i \rightarrow \infty$.²

Etter hvert som vi har konvergens vil vi ha $e_{i+1} \approx S / e_i$.

¹Dette er teorem 1.6 i læreboka.

²Her har vi **utelatt noen detaljer i beviset** for teoremet, se SAUER s. 36. (Kreves ikke til eksamen.)

Oppgave 26

Anta at $g(x)$ er kontinuerlig deriverbar på et lukket intervall der $a \leq g(x) \leq b$ for $x \in [a, b]$ og $|g'(x)| < 1$.

Vis at da har $g(x)$ et og bare et fikspunkt i intervallet.

¹Da $|\alpha - \beta| \neq 0$, gir $|\alpha - \beta| < |\alpha - \beta|$, dvs. den opplagte selvmotsigelsen $1 < 1$.

Oppgave 26

Anta at $g(x)$ er kontinuerlig deriverbar på et lukket intervall der $a \leq g(x) \leq b$ for $x \in [a, b]$ og $|g'(x)| < 1$.

Vis at da har $g(x)$ et og bare et fikspunkt i intervallet.

Løsning:

Med $f(x) = x - g(x)$ er $f(a) \leq 0$ og $f(b) \geq 0$, så da sier skjæringssetningen at det er minst et fikspunkt for $g(x)$ i $[a, b]$.

¹Da $|\alpha - \beta| \neq 0$, gir $|\alpha - \beta| < |\alpha - \beta|$, dvs. den opplagte selvmotsigelsen $1 < 1$.

Oppgave 26

Anta at $g(x)$ er kontinuerlig deriverbar på et lukket intervall der $a \leq g(x) \leq b$ for $x \in [a, b]$ og $|g'(x)| < 1$.

Vis at da har $g(x)$ et og bare et fikspunkt i intervallet.

Løsning:

Med $f(x) = x - g(x)$ er $f(a) \leq 0$ og $f(b) \geq 0$, så da sier skjæringssetningen at det er minst et fikspunkt for $g(x)$ i $[a, b]$.

La $\alpha = g(\alpha)$ og $\beta = g(\beta)$ være to fikspunkter $\alpha \neq \beta$ i $[a, b]$.

Da gir middelverdisetningen at $\alpha - \beta = g(\alpha) - g(\beta) = |g'(c)|(\alpha - \beta)$ for c i (a, b) , dvs. at $|\alpha - \beta| = |g'(c)||\alpha - \beta| < |\alpha - \beta|$ da $|g'(c)| < 1$, så $\beta = \alpha$, som påstått.¹

¹Da $|\alpha - \beta| \neq 0$, gir $|\alpha - \beta| < |\alpha - \beta|$, dvs. den opplagte selvmotsigelsen $1 < 1$.

Oppgave 26

Anta at $g(x)$ er kontinuerlig deriverbar på et lukket intervall der $a \leq g(x) \leq b$ for $x \in [a, b]$ og $|g'(x)| < 1$.

Vis at da har $g(x)$ et og bare et fikspunkt i intervallet.

Løsning:

Med $f(x) = x - g(x)$ er $f(a) \leq 0$ og $f(b) \geq 0$, så da sier skjæringssetningen at det er minst et fikspunkt for $g(x)$ i $[a, b]$.

La $\alpha = g(\alpha)$ og $\beta = g(\beta)$ være to fikspunkter $\alpha \neq \beta$ i $[a, b]$.

Da gir middelverdisetningen at $\alpha - \beta = g(\alpha) - g(\beta) = |g'(c)|(\alpha - \beta)$ for c i (a, b) , dvs. at $|\alpha - \beta| = |g'(c)||\alpha - \beta| < |\alpha - \beta|$ da $|g'(c)| < 1$, så $\beta = \alpha$, som påstått.¹

MERK:

Hvorfor konvergerer da FPI metoden med $g(x)$ som oppfyller betingelsene ovenfor for alle startverdier $x_0 \in [a, b]$?

¹Da $|\alpha - \beta| \neq 0$, gir $|\alpha - \beta| < |\alpha - \beta|$, dvs. den opplagte selvmotsigelsen $1 < 1$.

Vi skal bestemme en effektiv metode for å beregne kvadratrøtter.

La $s = \sqrt{a}$ for $a > 0$.

Hvis a_n er en tilnærmelse for s , da ligger s mellom a_n og a/a_n (hvorfor?).

¹Da $\sqrt{a} \in [a_n, a/a_n]$ eller $\sqrt{a} \in [a/a_n, a_n]$ (hvorfor?), er middelverdien også innenfor intervallet.

Vi skal bestemme en effektiv metode for å beregne kvadratrøtter.

La $s = \sqrt{a}$ for $a > 0$.

Hvis a_n er en tilnærmelse for s , da ligger s mellom a_n og a/a_n (hvorfor?).

Da vil middelverdien av disse to tilnærmelsene forhåpentligvis være en enda bedre tilnærmelse til s .

For en startverdi a_0 gir dette gir opphav til fikspunktiterasjonen

$$a_{n+1} = g(a_n) = (a_n + a/a_n)/2^1$$

¹Da $\sqrt{a} \in [a_n, a/a_n]$ eller $\sqrt{a} \in [a/a_n, a_n]$ (hvorfor?), er middelverdien også innenfor intervallet.

Vi skal bestemme en effektiv metode for å beregne kvadratrøtter.

La $s = \sqrt{a}$ for $a > 0$.

Hvis a_n er en tilnærmelse for s , da ligger s mellom a_n og a/a_n (hvorfor?).

Da vil middelverdien av disse to tilnærmelsene forhåpentligvis være en enda bedre tilnærmelse til s .

For en startverdi a_0 gir dette gir opphav til fikspunktiterasjonen

$$a_{n+1} = g(a_n) = (a_n + a/a_n)/2^1$$

Iterasjonen er svært effektiv, se nedenfor. (Vis at $g'(\sqrt{a}) = 0$!)¹

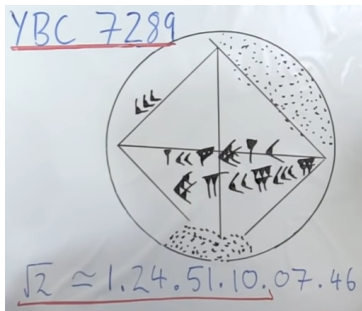
- **1** (a_0)
- $\frac{3}{2} = \mathbf{1.5}$ (a_1)
- $\frac{17}{12} = \mathbf{1.416...}$ (a_2)
- $\frac{577}{408} = \mathbf{1.414215...}$ (a_3)
- $\frac{665857}{470832} = \mathbf{1.4142135623746...}$ (a_4)

¹Da $\sqrt{a} \in [a_n, a/a_n]$ eller $\sqrt{a} \in [a/a_n, a_n]$ (hvorfor?), er middelverdien også innenfor intervallet.

¹Når vi behandler Newton iterasjon, skal vi se at dette sikrer såkalt **kvadratisk** konvergens.

Babylonerne brukte 60-talls-systemet og ga tilnærmelsen
1.24.51.10 for $\sqrt{2}$ med metoden beskrevet på forrige slide.¹

Det er virkelig imponerende at dette skjedde for så lang tid tilbake!¹



```
In[1]:= 1+24/60+51/60^2+10/60^3//Together
      30547
Out[1]= ----
      21600

In[2]:= N[%,10]
Out[2]= 1.414212963

In[3]:= 1+24/60+51/60^2+10/60^3+7/60^4+46/60^5//Together
      549846233
Out[3]= ----
      388800000

In[4]:= N[%,10]
Out[4]= 1.414213562

In[5]:= 2//Sqrt
Out[5]= Sqrt[2]

In[6]:= N[%,10]
Out[6]= 1.414213562
```

¹ Dette er kun for informasjon og kreves ikke til eksamen.

¹ For over 3500 år siden.

I stoppekriteriet nedenfor er $\theta > 0$.

- Trenger kriterium for å stoppe:
- Gitt toleranse: TOL
- Stopp kriterium:
 - ▶ Absolutt: Stopp når $|x_{i+1} - x_i| < TOL$.
 - ▶ Relativ: Stopp når $\frac{|x_{i+1} - x_i|}{|x_{i+1}|} < TOL$.
 - ▶ Hybrid: Stopp når $\frac{|x_{i+1} - x_i|}{\max(|x_{i+1}|, \theta)} < TOL$.