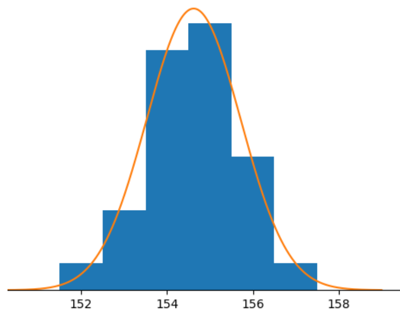
4 - 6 - OM DØDSLENGSEL

Hvis du skal bygge kai og har bestilt sekstoms bord, får du såkalte *råbord*. De er bra trykkimpregnert, men ikke saga så nøye, så de fleste av dem er ikke nøyaktig 152.4 mm, men kanskje et eller annet sted mellom 152mm og 157mm. Jeg fiksa kaien til svigers en sommer, og målte den faktiske bredden på alle de tjueni sekstomsbordene som var kjøpt inn. Gjennomsnittsbredden var $\hat{\mu} = 154.6$ mm, og det empiriske standardavviket på $\hat{\sigma} = 1.1$ mm. Til høyre er et plot av normalfordelingsfunksjonen oppå relative frekvenser, altså antall bord av hver lengde delt på totalt antall bord. Du finner dette som .csv-fil her: `https://folk.ntnu.no/mortano/python/kai`.

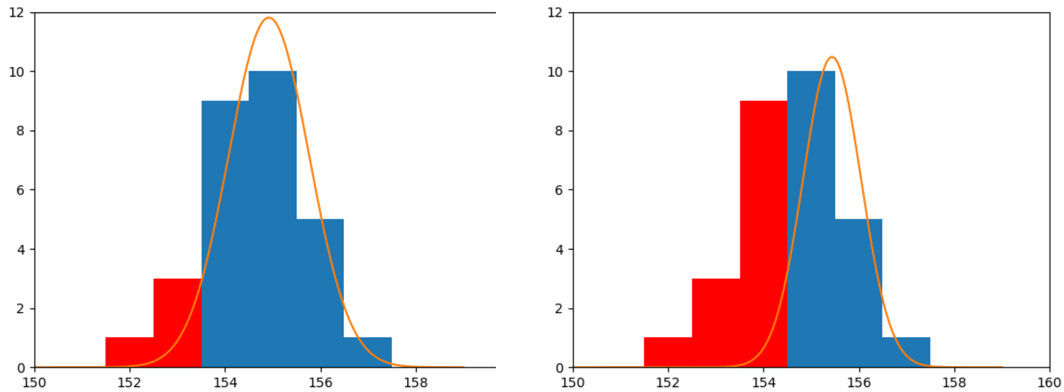


Nå skal vi komplisere det litt. Forestill deg at du har et måleapparat og at du måler en normalfordelt størrelse, men så vet du at den ene halen i normalfordelingen er under deteksjonsgrensen for måleapparatet. Hvis du ønsker å estimere μ og σ, kan du ikke ta gjennomsnitt og empirisk standardavvik, for gjennomsnittet til treffe for høyt og det empiriske standardavviket for lavt. I datasettet over bredden på kaibordene har jeg tatt ut de laveste målingene (tre bord på 153 mm og ett på 152 mm). Da får man datasettet i margen. Hvilken normalfordelingskurve

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

kommer dette fra? I denne økten skal vi prøve å finne det ut.
I det venstre plottet under finner du histogrammet til kaibordene og en normalfordelingskurve der μ og σ er estimert ved gjennomsnitt og empirisk standardavvik på data uten 152- og 153-målingene. Det er kanskje ikke så lett å se med det blotte øye at kurven er feil, så jeg har for illustrasjonens skyld tatt ut 154mm-målingene også og gjentatt eksperimentet; dette er plottet til høyre. I dette plottet er det ganske lett å se at noe er galt.

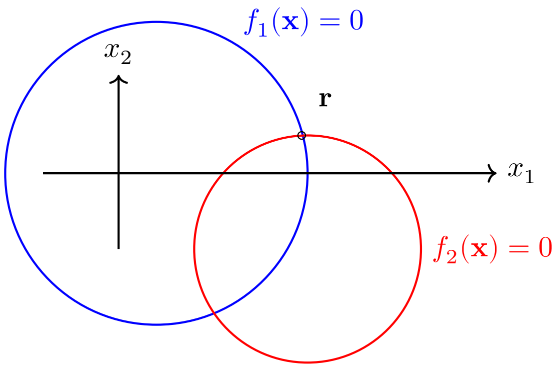
1	154
2	154
3	
4	156
5	155
6	
7	154
8	155
9	
10	154
11	155
12	155
13	154
14	156
15	155
16	154
17	155
18	154
19	154
20	156
21	155
22	154
23	155
24	157
25	
26	155
27	156
28	155
29	156



Problemet på forrige side er litt komplisert, så la oss begynne med noe mye enklere - Newtons metode i to dimensjoner. La oss si at vi ønsker å løse det ikkelineære likningssettet

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{pmatrix} = \mathbf{0}$$

der $\mathbf{f}$ er en funksjon fra $\mathbb{R}^2$ til $\mathbb{R}^2$. Dette er to likninger for nullnivåkurvene til komponentene til $\mathbf{f}$, som begge er funksjoner fra $\mathbb{R}^2$ til $\mathbb{R}$. En løsning $\mathbf{r}$ av likningssettet må altså ligge på skjæringspunktet mellom disse nivåkurvene. Her er en figur som illustrerer dette med to sirkulære nivåkurver:



Numeriske metoder er stort sett iterative. Vi lar $\mathbf{x}_n$ være iterasjon nummer n.

$\boxed{1}$ Sett opp tangentplanene til f_1 og f_2 i iterasjon $\mathbf{x}_n$, og utled **Newtons metode**

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{f}'(\mathbf{x}_n)^{-1}\mathbf{f}(\mathbf{x}_n)$$

(Hint: I bunn og grunn samme som den endimensjonale varianten vi hadde i TMA4101.)

$\boxed{2}$ Test ved å finne skjæringspunktene mellom de to ellipsene

$$3x^2 + 2xy + y^2 - 2x - y = 1 \qquad \text{og} \qquad x^2 + 2xy + 6y^2 - x - 2y = 1$$

Hint: Det går greit å bruke np.linalg.solve, men for 2×2-matriser finnes det en formel:

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

Apropos - **kofaktorekspansjon**. Det finnes en generell formel for inversen til en inverterbar matrise:
`https://en.wikipedia.org/wiki/Invertible_matrix#Analytic_solution`
Det er en og annen gubbe (inkludert meg) som sporadisk synes dette er nyttig å vite om. [1]

$\boxed{3}$ Forklar hvorfor det å løse oppgave 2 er det samme som å lete etter kritiske punkter til funksjonen

$$V(x, y) = x^3 + x^2y + xy^2 + 2y^3 - x^2 - y^2 - xy - x - y.$$

(Hint: Se på gradienten til V.)

[1] Jeg har hatt bruk for det én gang i livet - i den eneste artikkelen jeg har ført i pennen noen gang. Den var kjempebra og biveilederen min fikk den publisert i en obskur bulgarsk konferansejournal rett før han ble pensjonist.

En kjent og kjær teknikk for å lete etter kritiske punkter heter *også* **Newtons metode**.

$\boxed{4}$ Forklar hvorfor denne er gitt ved iterasjonen

$$\mathbf{x}_{n+1} = \mathbf{x}_n - V''(\mathbf{x}_n)^{-1}V'(\mathbf{x}_n)$$

der $V : \mathbb{R}^2 \to \mathbb{R}$ er funksjonen vi leter etter kritiske punkter til.
(Husk at V' er gradienten og V'' er hessematrisen.)

Antagelig oppdaget du at dødslengselmetoden konvergerer ganske sakte, til tross for at den prøver å gå rett utfor så bratt som mulig. Dette skyldes at den bratteste retningen ikke nødvendigvis peker mot det kritiske punktet. La oss se litt på dette og samtidig sette lineæralgebraen i nytt lys.

$\boxed{8}$ Forklar at for alle lineære systemer $A\mathbf{x} = \mathbf{b}$ du har løst til nå, er $\mathbf{x}$ minimumspunktet til

$$V(\mathbf{x}) = (A\mathbf{x} - \mathbf{b})^T(A\mathbf{x} - \mathbf{b}) = \|A\mathbf{x} - \mathbf{b}\|^2$$

I praktiske anvendelser må man ofte løse enorme lineære systemer der de fleste komponentene i matrisen er null. Da er det sjelden effektivt å gausseliminere. Når du bruker np.linalg.solve, spør python bare LAPACK[2] og der finnes det et tonn med forskjellige metoder. Vektoren $A\mathbf{x}-\mathbf{b}$ kalles **residualet**, og noen av de mest kjente iterative teknikkene for lineære systemer (for eksempel GMRES[3]) baserer seg på å minimere denne i forskjellige underrom av $\mathbb{R}^n$.

$\boxed{9}$ I anvendelser er matrsien A ofte symmetrisk og positivt definitt. Da kan vi minimere en enklere funksjon. Forklar at dersom $A \in \mathbb{R}^n$ er symmetrisk og positivt definitt, er det å løse $A\mathbf{x} = \mathbf{b}$ er det samme som å minimere funksjonen

$$V(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{x}^T\mathbf{b}.$$

Kvadratiske polynomer er praktiske å jobbe med i optimering, for de er lette å analysere; mange klassiske linjesøksmetoder baserer seg på å minimere andre ordens taylorpolynomer til den funksjonen som egentlig skal minimeres. Hvis V er kvadratisk slipper for eksempel backtracking.

$\boxed{10}$ Du kjører et steg med en linjesøksmetode på problemet over, og har valgt ut retningen $\mathbf{p}_n$. Finn minimumspunktet til

$$V(\mathbf{x}_n + \alpha_n \mathbf{p}_n)$$

som funksjon av α_n.

Nå kan vi se litt på hvorfor dødslengsel konvergerer sakte. Dersom A er positivt definitt, er nivåkurvene til f i oppgave 9 alltid ellipser.

$\boxed{11}$ Tegn opp noen litt lange og flate elliptiske nivåkurver og tegn noen dødslengseliterasjoner omtrentlig for hånd med α_n bestemt slik som som i oppgave 10.

[2]`https://en.wikipedia.org/wiki/LAPACK`
[3]`https://en.wikipedia.org/wiki/Generalized_minimal_residual_method`

En av de mest kjente metodene for dette kalles **konjugerte gradienter**, og denne er så bra at oppfinnerne Hestenes og Stiefel ble kjendiser. Den konvergerer som et uvær i mange tilfeller, men som de fleste numeriske metoder vil den også feile spektakulært i andre. Hvis du går MTKJ og treffer på Svein Sunde kan du spørre ham om denne metoden, han har brukt den i atomkraftverk.

$\boxed{13}$ Vis at

$$(\mathbf{x}, \mathbf{y})_A = \mathbf{x}^T A \mathbf{y}$$

er et indreprodukt dersom A er symmetrisk og positivt definitt.

Dersom $\mathbf{x}$ og $\mathbf{y}$ er ortogonale med hensyn på indreproduktet over, sier vi at de er A-ortogonale. Et velkjent resultat fra abstrakt lineæralgebra er at om du har et vektorrom og et indreprodukt, finnes det alltid en ortogonal basis.

$\boxed{14}$ Anta at du har en A-ortogonal basis $\{\mathbf{p}_k\}$ for $\mathbb{R}^n$, at $A\mathbf{x} = \mathbf{b}$, og at

$$\mathbf{x} = \sum_k \alpha_k \mathbf{p}_k.$$

Utled at

$$\alpha_k = \frac{(\mathbf{b}, \mathbf{p}_k)}{(\mathbf{p}_k, \mathbf{p}_k)_A}$$

Konjugerte gradienters metode består i å konstruere et sett med søksretninger som er A-ortogonale. Det viser seg at man på magisk vis får dette til ved rekursjonen

$$\mathbf{p}_{k+1} = -\mathbf{r}_{k+1} + \beta_{k+1} \mathbf{p}_k$$

der

$$\mathbf{r}_k = A\mathbf{x}_k - \mathbf{b}$$

er residualvektoren etter iterasjon k, og β_k velges slik at $\mathbf{p}_k$ og $\mathbf{p}_{k+1}$ er A-ortogonale.

$\boxed{15}$ Vis at

$$\beta_{k+1} = \frac{(\mathbf{r}_{k+1}, \mathbf{p}_k)_A}{(\mathbf{p}_k, \mathbf{p}_k)_A}.$$

Steglengden velges slik som i oppgave 10 for at det skal funke. Du finner full algoritme på s 108 i Nocedal og Wright.

$\boxed{16}$ Vis at α_k over minimerer V slik som i oppgave 10.

$\boxed{17}$ Prøv konjugerte gradienters metode på implisitt skjema for varmelikningen.

UKENS NØTTER

$\boxed{1}$ Estimér μ og σ i kaieksemplet på side 1. Dette kan gjøres på mange måter.

$\boxed{2}$ Har det noen hensikt å kjære Newtons metode på et lineært system?

Dersom A er symmetrisk og positivt definitt, finnes det en mye mer fornuftig retning å kjøre enn brattest utfor, nemlig langs med halvaksene til de elliptiske nivåkurvene.

$\boxed{3}$ Forklar at med strategien fra oppgave 10 trenger du egentlig bare to iterasjoner for å komme frem til minimumspunktet dersom du velger $\mathbf{p}_n$ slik.