

AST2000 Del 1C

Interaktive forelesningsnotater: forelesning 2 av 2

VIKTIG

Dette er et alternativ til forelesningen i emnet. Har du gått skikkelig gjennom disse interaktive forelesningsnotatene så trenger du ikke å lese de fulle forelesningsnotatene (med unntak av oppgavene bak). All informasjonen du trenger, får du her. Du kommer til å få mange grublespørsmål og diskusjonsoppgaver, det er meningen at disse skal gjøres i grupper av minst 2, maks 4 studenter. Det er defor sterkt anbefalt at dere sitter sammen i grupper når dere går gjennom disse interaktive forelesningsnotatene, du vil få betydelig mer utbytte av dem på den måten. Hvis du har kommentarer ris/ros til disse forelesningsnotatene eller til emnet, trykk på ☺ ☹ knappen som du finner på alle sider.

Trykk denne knappen for å begynne

- HUSK at du får mer ut av de interaktive forelesningsnotatene når du gjør de sammen med noen. Diskusjonene med andre er svært viktige.
- Det er mange spørsmål/grubliser underveis, sett dere selv en tidsgrense, 1 minutt på de korte, maks 4-5 minutter på de lengre. Ha en alarm ved siden av, ellers kommer dere til å bruke alt for langt tid. Har dere ikke fått det til etter kort tid, gå videre, se svaret og lær!
- Er du i det minste tvil om noe, så finnes det en **FORUM** knapp, trykk det og still spørsmål med en gang mens du enda husker spørsmålet!

Trykk denne knappen for å begynne

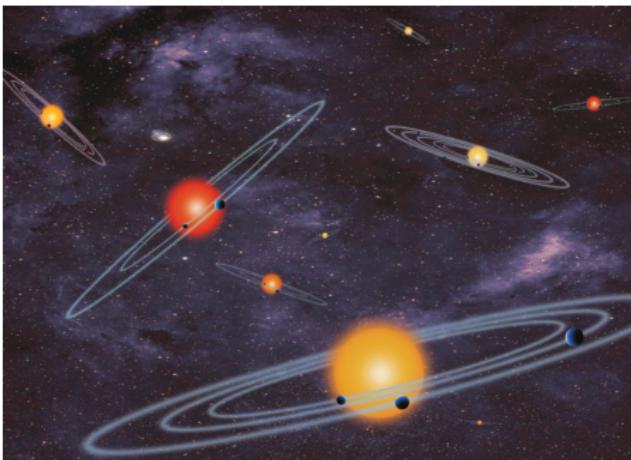
Forrige side

Hvis du allerede har begynt på denne forelesningen og vil hoppe rett inn til et annet kapittel, kan du trykke her:

- Et eksempel
- Støy i data: Teoretisk modell
- Minste kvadraters metode
- Støy i data:statistikk
- Tips til parameterestimering

Neste side

[Forrige side](#)



Velkommen til andre forelesning av del 1C! Vi skal begynne med å bruke det du lærte i første forelesning på noen eksempler. Deretter skal vi se hvordan vi kan analysere data med støy, og du skal lære å estimere planetmasser fra mer realistiske stjernedata. Dette forelesningsnotatet tilsvarer omtrent en dobbelttime fysisk forelesning.

Er du klar?

[Neste side](#)

Vi utledet i forrige forelesning sammenhengen mellom stjernas banehastighet og massen til planeten som går i bane rundt stjerna

$$m_p = \left(\frac{P}{2\pi G} \right)^{1/3} m_*^{2/3} \frac{v_{*r}}{\sin i}$$

der P er omløpsperioden, i er inklinasjonen (**husker du hvordan den er definert? Hvis ikke, gå tilbake og sjekk nå!**), stjernas masse m_* som vi kan regne som kjent og til slutt v_{*r} som er stjernas observerte radielle komponent av banehastighet i banen om massesenteret. Dopplereffekten gjør at vi kan finne stjernas radialhastighet (**sjekk at du husker definisjon av radialhastighet og tangensialhastighet fra observatøren**) på forskjellige tidspunkter ved å se på bølgelegnden til en spektrallinje. Fra hastighetskurven til stjerna (radialhastheten $v_r(t)$ som funksjon av tiden) kan vi finne banehastigheten v_* til stjerna (**sjekk at du husker hvordan du leser den av fra kurven**).

Inklinasjonen i er vanskelig å finne, men ved $i = 90^\circ$ kan vi få en formørkelse av stjerna hver gang planeten går foran. Deg gjør at vi kan måle radiusen til planeten hvis vi har gode data ved at $2R_p = v_p(t_1 - t_0)$. **Husker du hvilke tidspunkter t_0 og t_1 du må lese av og hvorfor disse tidspunktene? Du bør kunne utlede denne sammenhengen!**

Forrige side

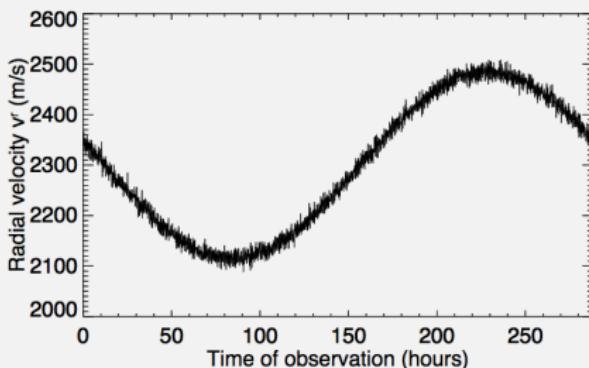
Nytt tema:

Et eksempel

Dette temaet fortsetter frem til side 11 av 42.

Sett igang med eksemplet!

Vi skal gjøre deler av oppgave 1C.3 her. Vi skal se på en realistisk hastighetskurve med støy. Vi skal snart lære mer om støy, men den gir et tilfeldig ekstra utslag på hastighetskurven i hvert tidssted. **Støyen gir et tilfeldig hastighetsbidrag i hvert tidssteg som i middel skal gå like mye opp som ned slik at den underliggende kurven uten støy ligger midt inne i støyen.**



Her ser vi radialhastighetskurven til en stjerne.

Anslå egenhastigheten til stjerna **hvis du ikke husker hva egenhastighet var, gå tilbake til forrige forelesning, dette er et svært viktig konsept!**. Er den:

0m/s

2500m/s

2475m/s

2300m/s

2125m/s

2100m/s

400m/s

175m/s

Forrige side



side 3 av 42

Et eksempel

FORUM

Det ble galt. Er du sikker på at du forstår hva egenhastighet er for noe? Egenhastigheten er hastigheten til massesenteret. Stjerna har hastigheten til massesenteret pluss banehastighet omkring massesenteret. Gå tilbake til forrige forelesning og kontroller at du forstår godt hvorfor hastighetskurven ser ut slik som den ser ut. Hvis du ikke forstår dette må du ikke gå videre da vil du ikke få noe ut av denne forelesningen. Tenk og gjør et nytt forsøk. Hvis du etter det enda ikke forstår, kikk på [denne videoen](#).

[Forrige side](#)



side 3 av 42

[Et eksempel](#)

[FORUM](#)

Bra! Du ser ut til å ha forstått hastighetskurven og at egenhastigheten er massesenterets hastighet. Hvis du likevel skulle være litt i tvil, så kikk gjerne på [denne videoen](#) som forklarer det i litt detalj.

[Neste side](#)

Forrige side

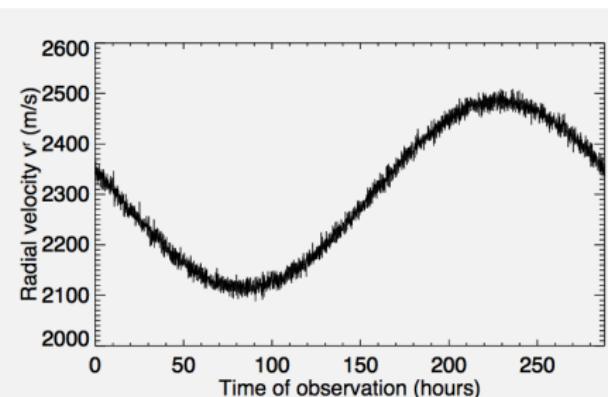


side 4 av 42

Et eksempel

FORUM

Vi fortsetter med deler av oppgave 1C.3 her:



Her ser vi radialhastighetskurven til en stjerne.

Hvis vi antar at inklinasjonen er 90° ,
anslå den fulle banehastigheten til
stjerna i bane om massesenteret. Er
den:

0m/s

2500m/s

2475m/s

2300m/s

2125m/s

2100m/s

400m/s

175m/s

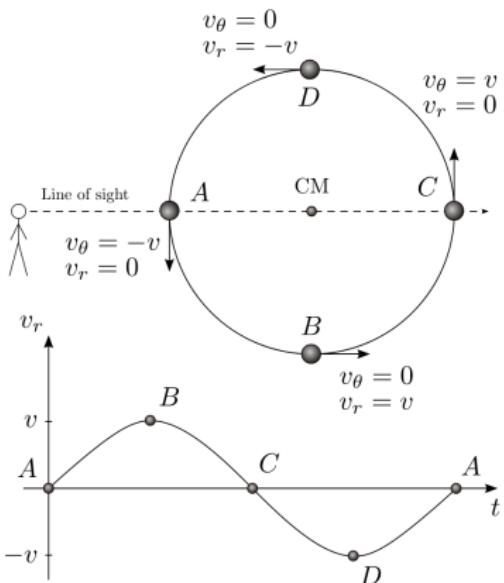
[Forrige side](#)

side 5 av 42

[Et eksempel](#)[FORUM](#)

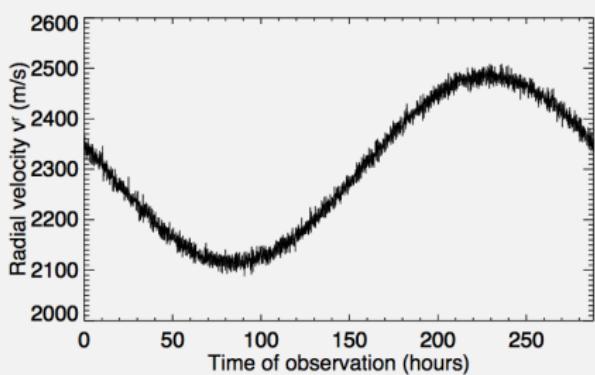
Det ble ikke helt riktig. Kan det være at du har glemt at støyen gir like mye utslag opp som ned og at den virkelige hastighetskurven som vi er ute etter ligger midt inne i den støyete kurven? Eller kan det være at du har glemt at banehastigheten er hastigheten **i forhold til massesenteret?** Tenk litt til, og prøv igjen!

Det er helt riktig! Siden banehastigheten er hastigheten i **forhold til massesenteret**, så må vi se på det maksimale utslaget fra massesenterhastigheten.



Vi ser på det maksimale utslaget siden v_r -komponenten har den fulle hastigheten i punkt B på figuren (se forrige forelesning). Massesenterhastigheten er 2300m/s og det maksimale utslaget på toppen av kurven er ca. 2475m/s. **MERK at det ikke er 2500 m/s siden dette er det maksimale utslaget med støy.** Hvis den virkelige glatte kurven ligger midt inne i støyen slik at det er like mye støyutslag opp som ned, så må den vel ligge på ca. 2475m/s? Dermed har vi at $2475\text{m/s} - 2300\text{m/s} = 175\text{m/s}$ (vi trekker altså fra massesenterhastigheten for å få hastighet i forhold til massesenteret).

Hvis vi nå antar at inklinasjonen er ukjent, gjør et estimat av den minste mulige massen som denne planeten må ha. Anta at man ved spektroskopi har anslått stjernas masse til å være $0.3m_{\odot}$ (merk at vi bruker m_{\odot} til å markere solmasser), altså en liten rød dvergstjerne. **Her trenger du formelen fra forelesning 1 som er repetert i starten her av forelesning 2**



Er den minste mulige massen:

- ca. 0.5 Jupitermasser
- ca. 0.9 Jupitermasser
- ca. 1.3 Jupitermasser
- ca. 2.1 Jupitermasser
- ca. 2.8 Jupitermasser
- ca. 3.5 Jupitermasser
- ca. 4.9 Jupitermasser

[Forrige side](#)

side 7 av 42

[Et eksempel](#)[FORUM](#)

Det ble nok ikke helt riktig. Usikkerheter i avlesninger gir litt slingringsmonn, men dette var litt for langt unna! Har du brukt $P = 300$ timer, $v_* = 175\text{m/s}$ og $m_* = 0.3m_{\odot}$? Hva får du hvis du bruker disse tallene? Blir det veldig forskjellig fra ditt tall? Gjør et nytt forsøk og spør gruppelærer hvis du ikke får det til.

[Forrige side](#)

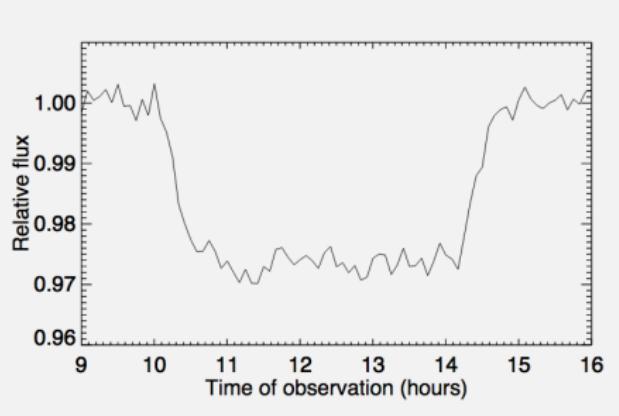
side 7 av 42

[Et eksempel](#)[FORUM](#)

Flott, du fikk det til! Kanskje du ikke fikk helt dette tallet? Det kan komme an på unøyaktigheter i avlesningen. Med $P = 300$ timer, $v_* = 175\text{m/s}$ og $m_* = 0.3m_{\odot}$ skulle du få et tall i nærheten av 0.9 jupitermasser.

[Neste side](#)

Som du ser blir det ganske høy usikkerhet da det er vanskelig å lese av P og v_* nøyaktig med støy på kurven. Dette skal vi se nærmere på om et par sider. Men først skal vi jobbe litt mer med oppgave 1C.3. Vi fortsetter med den samme stjerna, men observasjoner av lyskurven viser at planeten formørker stjerna. Her ser vi utdrag av lyskurven fra en formørkelse:



Gi et anslag på radiusen og tettheten til planeten. Du trenger kanskje å repetere fra forrige forelesning hvordan du finner v_p (i forhold til CM) når du har v_* .

Her blir det stor usikkerhet avhengig av hvordan du leser av den støyete kurven, svaret ditt ligger ikke nødvendigvis rett på et av disse, velg det nærmeste:

- ca. 0.5 Jupiterradier
- ca. 0.9 Jupiterradier
- ca. 1.3 Jupiterradier
- ca. 2.1 Jupiterradier
- ca. 2.8 Jupitradier
- ca. 3.5 Jupiterradier

[Forrige side](#)

side 9 av 42

[Et eksempel](#)[FORUM](#)

Det ble nok ikke helt riktig. Usikkerheter i avlesninger gir litt slingringsmonn, men dette var litt for langt unna! Har du brukt at tiden fra planeten begynner å formørke stjernen til planeten formørker stjernen fullstendig er ca. 0.8 timer? Og at planethastigheten $v_p \approx 60\text{km/s}$? Hva får du hvis du bruker disse tallene? Blir det veldig forskjellig fra ditt tall? Gjør et nytt forsøk og spør gruppelærer hvis du ikke får det til.

[Forrige side](#)

side 9 av 42

[Et eksempel](#)[FORUM](#)

Flott, du fikk det til! Kanskje du ikke fikk helt dette tallet? Det kan komme an på unøyaktigheter i avlesningen. Med $v_p \approx 60\text{km/s}$ og $t_1 - t_0 \approx 0.8$ timer skulle du få et tall i nærheten av 1.3 jupiterradier. Vi ser altså at vi har en planet som er litt mindre massiv en jupiter men med noe større radius. Da får vi ca. 0.6 ganger tettheten til vann, litt mindre enn Saturn som har 0.7 ganger vannets tetthet. Dette er altså helt klart en gassplanet.

$$\rho = \frac{0.9 \times 1.9 \times 10^{27}\text{kg}}{\frac{4}{3}\pi(1.3 \times 70 \times 10^6\text{m})^3}$$

[Neste side](#)

Forstår du alle stegene? Hvis ikke, ta kontakt med foreleser eller gruppelærer!

Husk at når vi karaktersetter prosjektet eller eksamen så er det forklaring av metoden som vi ser etter, ikke svaret. Når det gjelder estimering av planetmasser/radier, så er det fort gjort å venne seg til hvor man skal lese av på kurven og så sette disse tallene inn i en formel. Men da er det fort gjort å glemme **hvorfor** man skal lese av på den måten og hvorfor formelen ser ut som den gjør. Det er **disse tingene** du blir testes i, **ikke** om du kan lese av fra en kurve og sette inn i en formel, det forutsetter vi at du kan.

Da har tiden kommet til å teste forståelsen din litt mer i **dette skjemaet**. Det kan ta litt tid å gå gjennom, men **ikke** gå videre før du har sendt det inn da du trenger denne treningen i stoffet før du kan gå videre på statistikk og dataanalyse. **Hvis du sliter med noen av spørsmålene i skjemaet, så er det svært viktig at du spør foreleser eller gruppelærer da skjemaet tester sentral forståelse som du kommer til å trenge videre i kurset.**

Jeg har sendt inn skjemaet og er klar for dataanalyse-delen

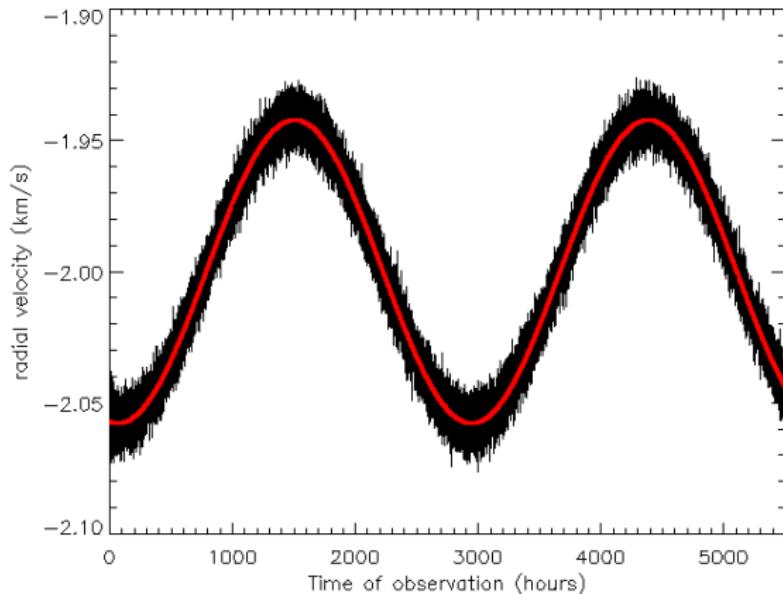
Forrige side

Nytt tema:

Støy i data: Teoretisk modell

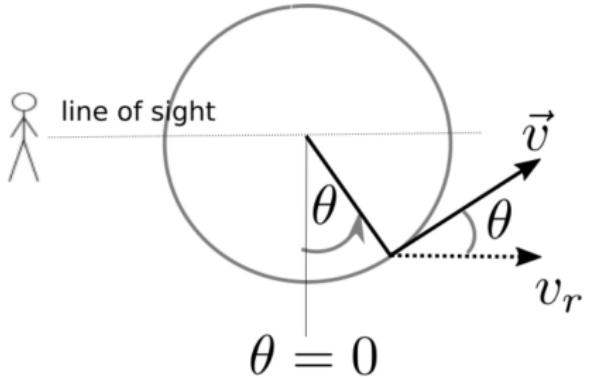
Dette temaet fortsetter frem til side 19 av 42.

Strekk litt på bena først og trykk her!



I figuren ser vi en radialhastighetskurve tatt fra en spektrallinje i spektret fra en stjerne. Den røde linjen viser hvordan kurven hadde sett ut hvis det ikke var støy i måleinstrumentene. Den sorte kurven viser hva som faktisk måles og som gjør det vanskelig å trekke nøyaktige hastigheter ut av kurven.

Støy i data: Teoretisk modell



Figuren viser stjerna i banen om massesenter med en vinkel θ som sier hvor i banen stjerna er. På figuren ser vi at θ også er vinkelen mellom \vec{v} og synslinja. Ser du hvorfor det er samme vinkel θ begge steder? (**hint:** en forlengelse av linja fra sentrum av sirkelen og radielt utover ved θ)..

For å kunne bruke kurven med støy til å finne mest mulig nøyaktige tall, trenger vi å lage en modell for den glatte 'virkelige' kurven som vi hadde observert hvis vi ikke hadde hatt støy. Bruk figuren til å se om du kan utlede en analytisk modell for observert radialhastighet som funksjon av vinkelen $\theta(t)$ som er vist på figuren, samt den faktiske banehastigheten v_* til banen (som er den vi er interessert i).

$$v_r(t) = v_* \sin \frac{2\pi}{\theta(t)}$$

$$v_r(t) = v_* \cos \frac{2\pi}{\theta(t)}$$

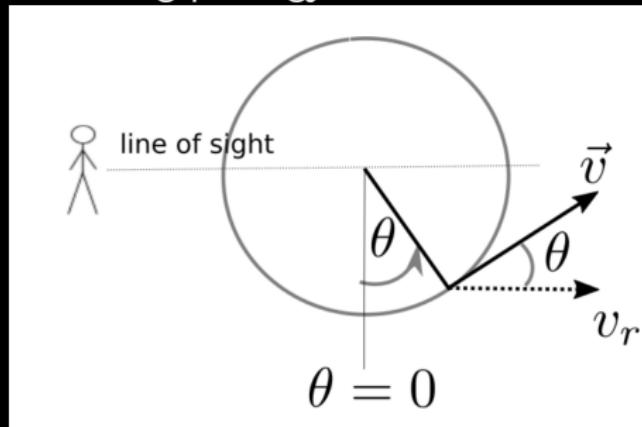
$$v_r(t) = v_* \sin \theta(t)$$

$$v_r(t) = v_* \cos \theta(t)$$

$$v_r(t) = v_* \cos \theta(t) \sin \theta$$

$$v_r(t) = v_* \tan \theta(t)$$

Det ble feil, kanskje du ikke kikket nøyde nok på figuren? Innser du hvorfor θ som er vinkelen mellom \vec{v} og v_r må være samme vinkel som den θ som angir hvor på sirkelen stjernen er? Hvis du gjør det så ser du at θ er vinkelen mellom \vec{v} og v_r , som betyr, som du jo vet, at v_r er x-komponenten av \vec{v} ? Dermed er v_r projeksjonen av \vec{v} med vinkelen θ . Hjelp det? Gå tilbake og prøv igjen.



[Forrige side](#)

side 13 av 42

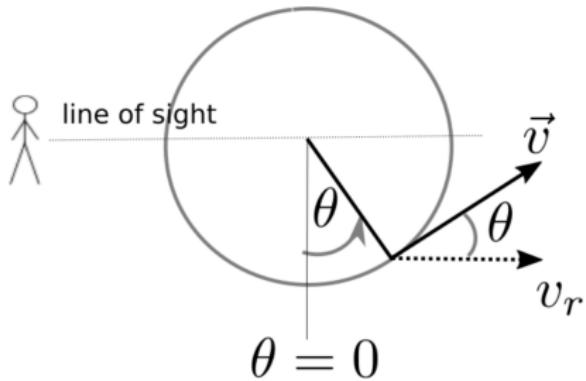
Støy i data: Teoretisk modell

[FORUM](#)

Helt riktig. Innser du hvorfor θ som er vinkelen mellom \vec{v} og v_r må være samme vinkel som den θ som angir hvor på sirkelen stjernen er? Hvis du gjør det, så har du innsett at v_r er x-komponenten til \vec{v} med vinkel θ imellom og dermed er altså

$$v_r(t) = v_* \cos \theta(t)$$

[Neste side](#)



Men for å ha en komplett modell, trenger vi også tidsavhengigheten til $\theta(t)$. Vi skal uttrykke den med omløpsperioden P . Anta at ved $t = 0$ så er $\theta = 0$. Er $\theta(t)$ lik:

$$\theta(t) = 2\pi Pt$$

$$\theta(t) = \frac{P}{t}$$

$$\theta(t) = P \cos 2\pi t$$

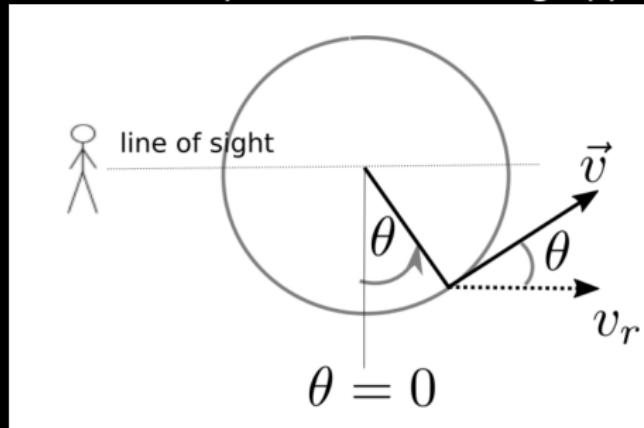
$$\theta(t) = \frac{2\pi P}{t}$$

$$\theta(t) = \frac{t}{P}$$

Vi ble enige om at

$$v_r(t) = v_* \cos \theta(t)$$

Njaaaa, det ble nok ikke helt riktig! Er du enig i at banehastigheten til stjerna er konstant siden det er sirkelbane? (hvis ikke, sjekk del 1B) Det betyr at vinkelhastigheten er konstant. Men hva er vinkelhastigheten? Stjerna skal gå rundt hele sirkelen i løpet av en periode P . Når du har vinkelhastigheten, altså endring i θ per tidsenhed, så kan du finne vi θ når du har den totale tiden. Hjelp det? Gå tilbake og prøv igjen. Sliter du med dette, spør foreleser eller gruppelærer.



Helt riktig. Stjerna skal gå en hel rundt på 2π i løpet av perioden P , dermed er vinkelhastigheten $\frac{2\pi}{P}$.

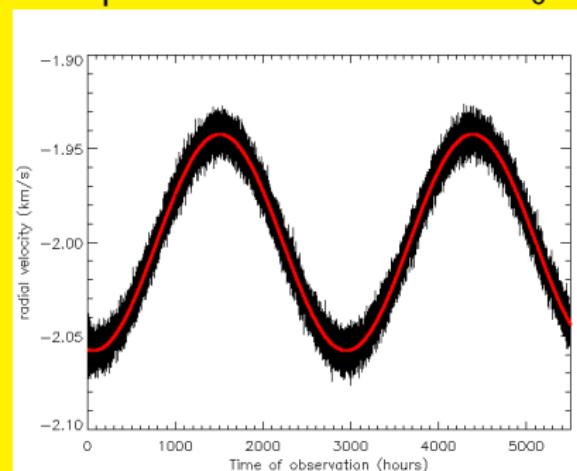
Ganger vi dette med tiden t som har gått så får vi vinkelen. Setter vi dette inn i vårt forrige resultat så har vi

$$v_r(t) = v_* \cos \frac{2\pi}{P} t$$

På figuren hadde vi definert når vinkelen θ er 0. Vi går nå bort ifra denne definisjonen og sier at ved tiden $t = t_0$ så er $\theta = 0$. Vi definerer altså isteden et tidspunkt t_0 der vinkelen er 0:

$$v_r(t) = v_* \cos \frac{2\pi}{P} (t - t_0)$$

Hvis vi definerer t_0 på denne måten, hvor på kurven er vi ved $t = t_0$??

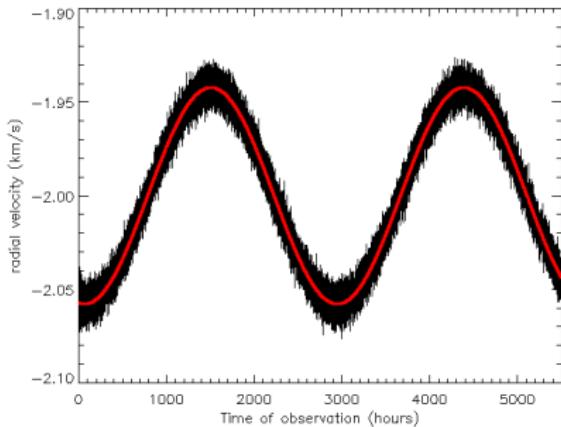


Tenk før du blir om (maks 1 min)! [Neste side](#)

Forrige side

**Er du helt sikker på at
du tenkte deg godt om nå???**
Ikke no' juks her! Hvis du
ikke hadde tenkt deg om, gå
tilbake til forrige side og
tenk! Hvis du hadde tenkt:
beklager så mye, det var
absolutt ikke meningen å
insinuere at du ikke følger
instrukser.

Neste side



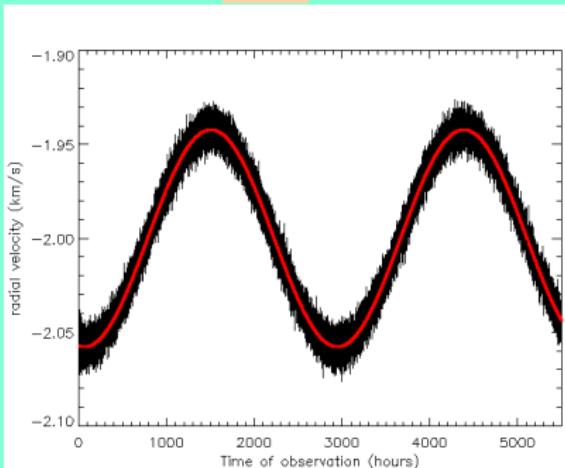
Hvis vi ser på uttrykket:

$$v_r(t) = v_* \cos \frac{2\pi}{P} (t - t_0)$$

så ser vi at ved $t = t_0$ så er $v_r = v_*$, altså radialhastigheten er lik den fulle banehastigheten.

Da er radialhastigheten maksimal, altså er vi på toppen av kurven. Dette skulle du huske fra diskusjonen i forrige forelesning: dette er punktet hvor stjernen er på vei bort fra oss langs vår synslinje og hele banehastigheten er i radialhastigheten uten noen tangensialkomponent (hvis vi antar at inklinasjonen er 90° !). Dette var punktet B på figuren i forrige forelesning. Repeter litt hvis du er usikker.

[Neste side](#)



$$v_r(t) = v_* \cos \frac{2\pi}{P} (t - t_0)$$

Vi har nå et uttrykk for hvordan kurven skal se ut hvis vi ikke har støy, uttrykt ved banehastigheten v_* , omløpsperioden P og tidspunktet t_0 når kurven har første topp-punkt. Problemet er bare at vi ikke kjenner noen av disse 3 parameterene.

Det er akkurat disse størrelsene vi ønsker å finne!!. Hvis vi finner v_* og P så kan vi estimere massen til planeten slik som i forrige forelesning. Men vi har alltså en støyete kurve som gjør det vanskelig å lese av v_* og P , samtidig har vi en analytisk modell som sier hvordan den underliggende støyfrie kurven skal se ut, dersom vi kjenner v_* og P . **Det store spørsmålet** nå er, hvordan kan vi bruke den analytiske modellen vår, sammen med den observerte støyete kurven til å finne gode estimater av tallene v_* og P ? Diskuter med medstudenter, og ikke gå videre før du har tenkt godt gjennom problemstillingen (3 min)!!!

Det var noen setninger på den forrige siden som er ganske så uforstålige hvis du ikke leser de flere ganger og tenker deg godt om. **Ta et skritt tilbake og se om du faktisk forstod hva som stod der, ok?** Når du har forstått hva som stod der, har du tenkt litt? Har du fått noen ideer? Tenk litt til før du går videre!

Forrige side

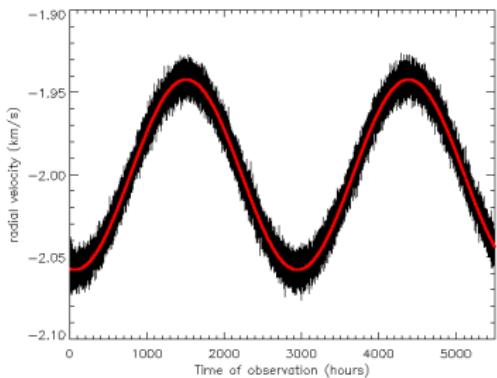
Nytt tema:

Minste kvadraters metode

Dette temaet fortsetter frem til side 22 av 42.

Vi tar kaffepause ved neste tema, litt til først!

Har du hørt om *minste kvadraters metode*? Metoden som går ut på å **minimalisere forskjellen mellom modellen og dataene?**



Vi skal finne den røde linja som passer best til de observerte dataene med støy.

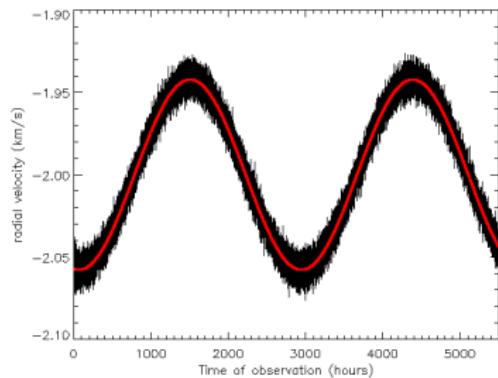
Vi vet at de røde linjene vi kan tegne, alle må følge:

$$v_r(t) = v_* \cos \frac{2\pi}{P} (t - t_0)$$

Vi kan endre på den røde linja ved å endre på verdiene for v_* , P og t_0 . **Minste kvadraters metode sier oss** altså å justere modellen (den røde linja), dvs. endre de ukjenten parameterene v_* , P og t_0 , helt til vi finner den modell-linja som gir minst forskjell mellom modellen og den observerte kurven.

Neste side

Minste kvadraters metode går altså ut på å **minimalisere forskjellen mellom modellen og dataene**:



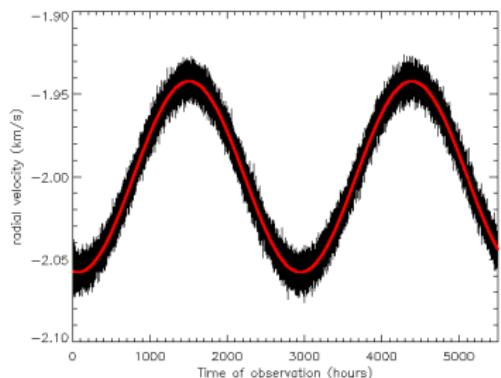
Der vi vet at modellen vår er gitt ved:

$$v_r(t) = v_* \cos \frac{2\pi}{P} (t - t_0)$$

Men hvordan defineres **den minste forskjellen?**. Det er rett og slett **forskjell** slik vi vanligvis omtaler en forskjell: **datakurven minus modellkurven**. Men vi opphøyer denne forskjellen i annen potens, derav navnet **minste kvadrat**. Dette er for å få absoluttverdien, det er ikke viktig om forskjellen er negativ eller positiv. Det finnes også en statistisk begrunnelse for å ta kvadratet, denne kommer snart.

[Neste side](#)

Vi skal altså lage oss mange forskjellige røde kurver med forskjellige tallverdier for P , v_* og t_0



$$v_r^{\text{mod}}(t) = v_* \cos \frac{2\pi}{P}(t - t_0)$$

helt til vi finner en som gir oss den minste mulige kvadratforskjellen. Rent matematisk skal vi altså summere kvadratforskjellen over alle målepunktene våre, dvs over alle observerte tidspunkter t_i .

Tidspunktet for observasjon nr. i er derfor t_i og vi har totalt N slike observasjoner. Da får vi at kvadratforskjellen er gitt ved:

$$\Delta(v_*, P, t_0) = \sum_{i=1}^N (v_r^{\text{obs}}(t_i) - v_r^{\text{mod}}(t_i, v_*, P, t_0))^2$$

der vi ser at Δ altså er en funksjon av de ukjente parameterene P , v_* og t_0 . Her er v_r^{obs} den observerte sorte kurven og v_r^{mod} er den røde modellkurven. Vi skal altså finne de parameterene P , v_* og t_0 som minimaliserer Δ . Dette vil være de tallene P , v_* og t_0 som passer best til dataene og de vi skal bruke videre til å finne massen.

[Neste side](#)

Forrige side

TID FOR KAFFEPAUSE!!!



Ut å strekke på bena. Inn med litt koffein...

Og kanskje kan du tillate deg en sjokoladebit? Men ikke si det høyt!

Ihværtfall 10 min...

Jeg lover, jeg har tatt pause og er klart til å fortsette

Forrige side

Nytt tema:

Støy i data:statistikk

Dette temaet fortsetter frem til side 37 av 42.

Sett igang med nytt tema!

Men er dette virkelig det beste vi kan gjøre? Det blir jo noen unøytigheter i verdiene for P , v_* og t_0 som vi får med denne metoden. Hvor presise resultatene våre er avhenger av hvor kraftig støy det er. Hvis vi nå skulle kjenne noen egenskaper til støyen, noe vi normalt gjør, kan vi da finne en mer sofistikert metode som gir oss bedre estimater? For å finne ut av dette, la oss dykke litt ned i



statistikken her...

Neste side

Normalt kan vi anta Gaussisk støy. Dvs at vi kan skrive hastigheten v_r i observasjon nr. i som

$$v_r^{\text{obs}}(t_i) = v_r^{\text{real}}(t_i) + \delta v_i$$

der v_r^{real} er det faktiske hastighetskurven som du hadde sett hvis det ikke hadde vært støy i instrumentet og δv_i er støy i observasjon i . Hvis støyen er Gaussisk så betyr det at δv_i er et tilfeldig tall trukket fra en Gaussisk sannsynlighetsfordeling. Normalt er denne med middelverdi 0 slik at hvis man hadde en uendelig mengde observasjoner av det samme datapunktet så ville du i middel fått den faktiske verdien. I tillegg kan man ofte måle standardavviket til fordelingen, vi kaller denne σ_n der n står for noise. Gitt at denne er kjent, skriv ned sannsynlighetsfordelingen $P(\delta v_i)$, dvs. sannsynligheten for at vi ved en gitt observasjon får en verdi δv_i for støyen. Er du klar? Har du et forslag til sannsynlighetsfordeling? Er du usikker, ta gjerne en kikk tilbake til del 1A for å reptere litt om Gaussiske sannsynlighetsfordeling. Du kommer til å få bruk for dette igjen, så gjør det med en gang, så du er klar til å gå videre!.

[Forrige side](#)

side 24 av 42

Støy i data/statistikk

[FORUM](#)

I del 1A lært vi at en Gaussisk sannsynlighetsfordeling for en stokastisk variabel x med gjennomsnittsverdi μ og standardavvik σ er gitt ved

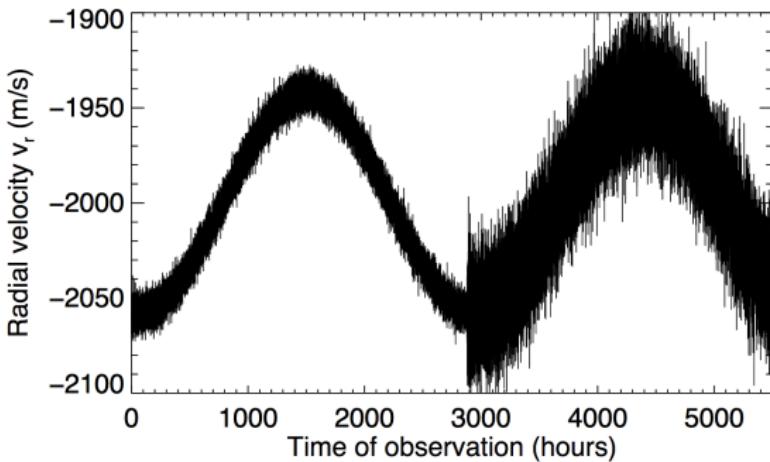
$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Hvis middelverdien for støyen er 0, noe den ofte er, og standardavviket er σ_n , så bli jo det da rett og slett:

$$P(\delta v_i) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{1}{2}\frac{\delta v_i^2}{\sigma_n^2}}$$

som igjen er en *sannsynlighetstetthet*, dvs at vi må gange med et intervall av mulige verdier for støyen δv_i for å få en faktisk sannsynlighet.

[Neste side](#)



På figuren ser du observasjoner av en hastighetskurve med to forskjellige teleskoper, et med betydelig høyere støynivå enn det andre. Dermed er standardavviket σ_n for støyen også forskjellig. Omtrent hva er σ_n for hver av disse instrumentene? La oss kalle observasjonene til venstre med minst støy for A og de til høyre med mest støy for B.

$$\sigma_n^A \approx 1930 \text{ m/s} \text{ og } \sigma_n^B \approx 1900 \text{ m/s}$$

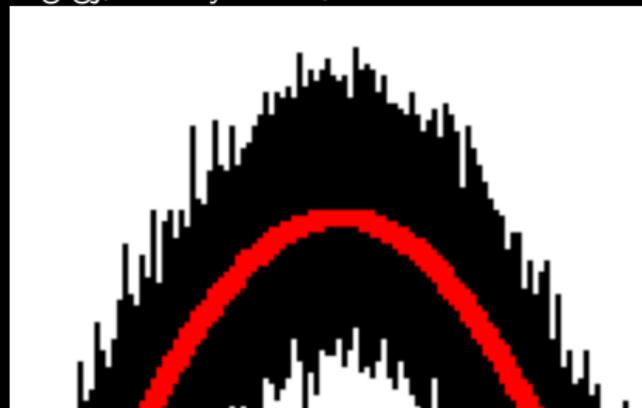
$$\sigma_n^A \approx 140 \text{ m/s} \text{ og } \sigma_n^B \approx 200 \text{ m/s}$$

$$\sigma_n^A \approx 30 \text{ m/s} \text{ og } \sigma_n^B \approx 80 \text{ m/s}$$

$$\sigma_n^A \approx 15 \text{ m/s} \text{ og } \sigma_n^B \approx 40 \text{ m/s}$$

$$\sigma_n^A \approx 7 \text{ m/s} \text{ og } \sigma_n^B \approx 20 \text{ m/s}$$

Det var ikke helt riktig. Under ser du en zoom av området rundt toppen av kurven til venstre (vi kunne ha tatt hvor som helst på kurven). Den røde linjen er observasjonen du hadde hatt uten støy, dvs. når støyen er 0. Utsalgene over og under er når du får positive eller negative verdier for støyen δv_i . Husker du at 95% av verdiene for en Gaussisk fordeling skal ligge innenfor 2σ ? **Hvis ikke, så er det svært viktig at du går tilbake til del 1A da du har glemt viktig kunnskap som du trenger videre.** Ta nå et papir, tegn en Gaussisk sannsynlighetsfordeling med forskjellige verdier for δv_i på x-aksen, med da 0 i midten siden gjennomsnittsverdien for denne Gaussiske fordelingen er 0. Tegn inn 1σ og 2σ (gjerne med hjelp av kurver fra del 1A) og sammenlikn med figuren under. Gå nå tilbake og gjør et nytt forsøk.



[Forrige side](#)



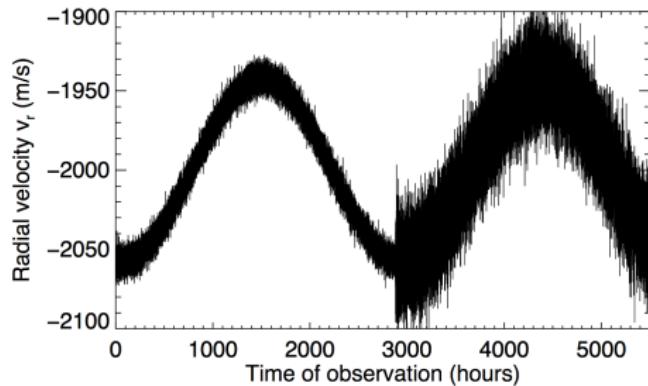
side 26 av 42

[Støy i data/statistikk](#)

[FORUM](#)

Flott! Du fikk det til. Det betyr at du forstår hva det betyr at støyen er Gaussisk fordelt og hva en Gaussisk fordeling er. Hvis du nå likevel er litt usikker, ta en titt på [denne videoen](#).

[Neste side](#)



Men hva skal vi gjøre i en slik situasjon når støyen varierer med tiden? Hvis vi bruker vanlig minste kvadraters metode her, så behandler vi dataene av dårlig kvalitet på samme måte som dataene av god kvalitet, kan det være riktig? Vi kan vel ikke få et optimalt godt estimat av modellen vår på den måten? Kan du tenkte deg hvordan vi kan gjøre en liten modifikasjon av minste kvadraters metode

$$\Delta(v_*, P, t_0) = \sum_{i=1}^N (v_r^{\text{obs}}(t_i) - v_r^{\text{mod}}(t_i, v_*, P, t_0))^2$$

for å ta hensyn til at standardavviket til støyen ikke alltid er den samme? **Tenk deg om (1 min) før du går videre!**

La oss se om vi kan utlede hva som er den beste måten å gjøre dette på!

La oss fortsette med statitsikken. Vi fant at sannsynligheten for støyen i observasjon i er gitt ved

$$P(\delta v_i) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{1}{2}\frac{\delta v_i^2}{\sigma_n^2}}$$

Hvis vi antar at standardavviket til støyen endrer seg for hver observasjon i og at standardavviket for støyen for observasjon nr. i er gitt ved σ_i , så blir vel den modifiserte sannsynligheten

$$P(\delta v_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\frac{\delta v_i^2}{\sigma_i^2}}$$

Enig?

Ja

Nei

[Forrige side](#)



side 29 av 42

[Støy i data:statistikk](#)

[FORUM](#)

Hvorfor ikke???? Er du usikker, spør foreleser!

[Neste side](#)

Flott, da kan vi gå videre. Vi har altså

$$P(\delta v_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2}\frac{\delta v_i^2}{\sigma_i^2}}$$

for sannsynligheten i et gitt tidspunkt i . Men hva er sannsynligheten for at vi får et helt sett med verdier $\{\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N\}$? Dvs. hva er f.eks. sannsynligheten for at vi får akkurat alle de støyverdiene som vi ser på figuren for alle tidspunkter i ? (**Hint: lov om multiplikasjon av sannsynlighet**)

Ikke gå til neste side før du har et forslag til sannsynlighetsfordeling for $P(\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N)$

Neste side

Netttopp ja, loven om multiplikasjon av sannsynligheter, vi ganger sammen sannsynligheten for støyen hvert tidspunkt i for å få den totale sannsynligheten for alle støyverdiene i alle N observasjonstidspunkter (hvis vi går tilbake til eksemplet vårt i del 1A så kan du jo tenke deg at den første støyverdien er høyden til en tilfeldig person, den andre er IQ-en til en tilfeldig person, den tredje er vekta til en tilfeldig person, etc. Ved å gange sammen sannsynlighetene kan du finne sannsynligheten for at en person har et gitt sett med verdier for disse 3 egenskapene, akkurat som vi nå kan finne sannsynligheten for at støyen i hvert observasjonspunkt har et gitt sett med verdier).

$$P(\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2} \frac{\delta v_i^2}{\sigma_i^2}}$$

[Forrige side](#)

side 31 av 42

[Støy i data/statistikk](#)[FORUM](#)

Ser du at vi kan skrive om på følgende måte...dette var uttrykket vår:

$$P(\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N)) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{1}{2} \frac{\delta v_i^2}{\sigma_i^2}}$$

og ved å skrive ut faktorene etter hverandre og bruke det vi kan om eksponenter, så kan vi skrive det som

$$P(\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N) = \left(\frac{1}{\sqrt{2\pi}} \right)^N \frac{1}{\prod_{i=1}^N \sigma_i} e^{-\frac{1}{2} \sum_{i=1}^N \frac{\delta v_i^2}{\sigma_i^2}}$$

Kan du se at det stemmer? Hvis ikke, skriv dette ut på papir og bruk loven om at man kan summe eksponentene i et produkt. **Ser du det?**

Hvis ikke spør foreleser![Neste side](#)

Hvis du går noen sider tilbake vil du se at vi skrev:

$$v_r^{\text{obs}}(t_i) = v_r^{\text{real}}(t_i) + \delta v_i$$

Dette skriver vi om til

$$\delta v_i = v_r^{\text{obs}}(t_i) - v_r^{\text{real}}(t_i)$$

og setter inn for δv_i i uttrykket vårt. Da får vi

$$P(\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N) = \left(\frac{1}{\sqrt{2\pi}} \right)^N \frac{1}{\prod_{i=1}^N \sigma_i} e^{-\frac{1}{2} \sum_{i=1}^N \frac{(v_r^{\text{obs}}(t_i) - v_r^{\text{real}}(t_i))^2}{\sigma_i^2}}$$

Målet vår er å finne en modell for $v_r^{\text{real}}(t_i)$, dvs. en modell $v_r^{\text{mod}}(t_i)$ som vi setter som vårt estimat av den reelle hastighetskurven. Hvis du kikker på uttrykket, begynner du å ane hvor minste kvadraters metode kommer fra?

Og begynner du ane hvordan den kan modifiseres?



Hvis vi setter inn modellen vår som den virkelige kurven så blir det:

$$P(\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N) = \left(\frac{1}{\sqrt{2\pi}} \right)^N \frac{1}{\prod_{i=1}^N \sigma_i} e^{-\frac{1}{2} \sum_{i=1}^N \frac{(v_r^{\text{obs}}(t_i) - v_r^{\text{mod}}(t_i))^2}{\sigma_i^2}}$$

hvor vi vet at $v_r^{\text{mod}}(t_i)$ inneholder de ukjente parameterene v_* , P og t_0 som vi ønsker å finne. Vi ønsker altså å finne den modellen, dvs. de parameterene v_* , P og t_0 , som gir den største mulige sannsynligheten for å observere den kurven, dvs. det sett med verdier $\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N$ som vi faktisk observerer. Vi ønsker altså å finne de verdiene for v_* , P og t_0 som maksimaliserer sannsynligheten $P(\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N)$. Hvordan kan vi skrive dette matematisk litt enklere??

[Neste side](#)

[Forrige side](#)

side 34 av 42

[Støy i data:statistikk](#)[FORUM](#)

Vi har

$$P(\delta v_0, \delta v_1, \delta v_2, \dots, \delta v_N) = \left(\frac{1}{\sqrt{2\pi}} \right)^N \frac{1}{\prod_{i=1}^N \sigma_i} e^{-\frac{1}{2} \sum_{i=1}^N \frac{(v_r^{\text{obs}}(t_i) - v_r^{\text{mod}}(t_i))^2}{\sigma_i^2}}$$

som skal maksimaliseres. Er du enig i at dette er identisk med å maksimalisere eksponenten her? Og siden det står et minustegn i eksponenten, er du enig i at vi da skal **minimalisere** det som kommer etter minustegnet? Tenk to ganger for å sjekke om du er enig!

[Neste side](#)

Er vi dermed enige om at for å maksimalisere sannsynligheten, så må vi **minimalisere**

$$\sum_{i=1}^N \frac{(v_r^{\text{obs}}(t_i) - v_r^{\text{mod}}(t_i))^2}{\sigma_i^2}$$

Denne størrelsen pleier vi å kalle χ^2 , dvs.

$$\chi^2 = \sum_{i=1}^N \frac{(v_r^{\text{obs}}(t_i) - v_r^{\text{mod}}(t_i))^2}{\sigma_i^2}$$

Hvis nå standardavviket til støyen σ_i er en konstant størrelse, hva blir dette til da???

Si det...

noe renspikka tull!

Newtons 2.lov!

Maxwell-Boltzmann!

Teorien om alt

Minste kvadraters metode

Hmmm, du har kanskje ikke fulgt så godt med?? Eller kanskje du falt av litt tidligere? Eller kanskje du var nysjerrig på hva som kom her? Du trenger nok å reptere litt av det grunnleggende som vi har gått gjennom før du kommer hit på nytt

[Trykk her!](#)

Flott! Helt riktig, tar vi ut σ_i som en konstant så får vi nøyaktig likningen for minste kvadraters metode som vi hadde tidligere. Vi har altså utledet minste kvadraters metode. Ser du nå også hvilken modifikasjon vi får når standardavviket til støyen endrer seg med observasjoner?

$$\chi^2 = \sum_{i=1}^N \frac{(v_r^{\text{obs}}(t_i) - v_r^{\text{mod}}(t_i))^2}{\sigma_i^2}$$

vi deler altså hvert kvadrat med σ_i^2 . Hvis støyen for denne målingen er stor, altså σ_i er stor, så deler vi på et stort tall. Omvendt, er støyen liten så deler vi på et lite tall. **Ser du at vi her vekter hvert kvadrat med et stort tall når kvaliteten på målingen er god, og et lite tall når kvaliteten er dårlig?.** Dette er χ^2 -metoden for estimering av parametere. **Vi antok Gaussisk støy i denne utledningen, hvis støyen ikke er Gaussisk, så er disse metodene ikke optimale for å estimere parametere!** Hvis avviket fra Gaussianitet ikke er veldig stort, så blir disse metodene likevel ofte brukt.

[Forrige side](#)

Vi er nesten i mål, men dette var en litt småøff utledning, ikke sant? Selv om det bare mangler et par sider, så trengs det nå litt tid til å fordøye stoffet.

Hvor mange ganger klare du å løpe opp og ned trappene til inngangen i Fysikkbygningen før du bare må legge deg ned og ikke klarer å løfte beina lenger? Sjekk hvem som klarer flest ganger. **Deretter, og kun deretter,** kan du gå videre... [Neste side](#)

Forrige side

Hvor mange ganger ble det???

1/2 gang: kun ned trappa

< 2

< 5

< 10

≥ 10

[Forrige side](#)

Ærlig talt! Skjerp deg!
Hardtrenings, hver dag fra nå av:
opp trappa 10 ganger om dagen!

[Neste side](#)

Forrige side

Seriøst? Prøvde du skikkelig? Tror jeg ikke, et forsøk til, la oss se om du klarer flere denne gangen!

Jepp, nå fikk jeg 50 ganger...

Forrige side

Det skulle bare mangle.
Men hvis formen er så god, hvorfor
prøver du ikke en gang til da,
kanskje du får til dobbelt så mange
denne gangen? Jepp, nå fikk jeg 50 ganger...

Forrige side

Når du har fått pusten tilbake, tenk nå godt gjennom:

- Hva er selve prinsippet/hovedideen bak minste kvadraters metode, og se for øyeblikket bort ifra den statistiske utledningen vi nettopp gjorde. Hva er essensen? Hvorfor funker minste kvadraters metode?
- Hva er forskjellen mellom minste kvadraters metode og χ^2 -minimalisering? Når trenger du χ^2 minimalisering og nøyaktig hva er forskjellen i det matematiske uttrykket? Igjen uten å tenke på den statistiske utledningen, hvordan og hvorfor funker χ^2 ?
- Så til den statistiske utledningen: Hva var utgangspunktet for utledningen? Og hva var konklusjonen, dvs. fra et rent statistisk ståsted hva er prinsippet bak minste kvadrat/ χ^2 ? Sett fra dette perspektivet, hvordan/hvorfor funker metodene?

Svarene kan du finne på de foregående sidene. Men før du leter der, se hvor mye du klarer å si om disse 3 punktene allerede nå. Etter det kan du nå bla tilbake og klare opp i disse 3 punktene! Hvis det er ting som er uklare/du er usikker på noe, kontakt foreleser!

Neste side

Forrige side

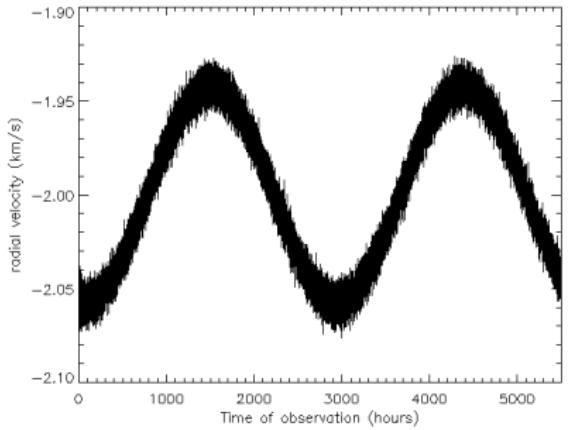
Nytt tema:

Tips til parameterestimering

Siste tema i denne forelesningen!

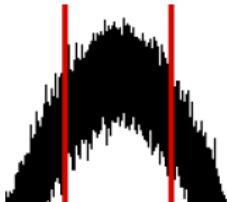
Okei, sett igang, nesten i mål!

Da har vi kommet til veis ende i dette kapittelet. Av god gammel vane, avslutter vi med noen tips til programmering og til slutt til debugging. I innleveringsoppgavene/prosjekt skal du analysere noen radialhastighetskurver med støy for å finne massen til en ekstrasolar planet. Da trenger du å bruke minste kvadraters metode eller χ^2 avhengig av om standardavviket til støyen varierer eller ikke. Tenk deg at du skal finne den beste mulige modellen (og dermed de beste mulige parameterene v_* , P og t_0) til følgende kurve:



Trykk her!

Du skal altså prøve deg frem med kombinasjoner av verdier for v_* , P og t_0 . Hver kombinasjon gir en modell $v_r^{\text{mod}}(t_i)$ som du så skal beregne kvadratet av forskjellen til de observerte dataene for. Sånn skal du holde på helt til du finner den kombinasjonen som gir minst mulig forskjell. Hvordan går vi frem? En mulighet er å finne et intervall av mulige parametere for hver av de 3 parameterene. Du kan f.eks. finne den verdien av t_0 som passer best ved å se i hvilket intervall denne verdien må ligge. På figuren

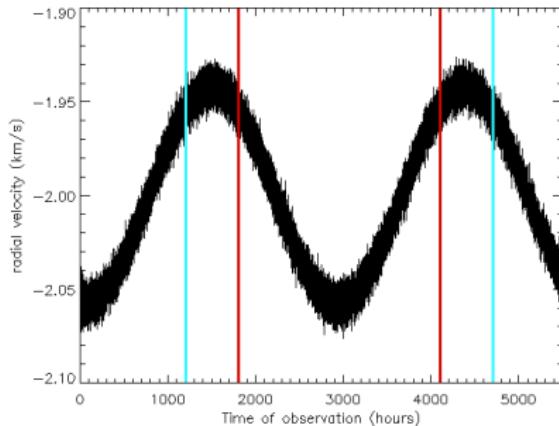


ser vi et mulig sånt intervall. Er du enig i at t_0 helt sikkert må ligge innenfor dette intervallet? Den kan vel ikke ligge utenfor? (hvis du har glemt hva t_0 står for, gå noen sider tilbake, det er viktig at du har dette helt klart for deg!). Del så dette intervallet opp i 20 like steg, og vips du

har 20 mulige verdier å prøve for t_0 .

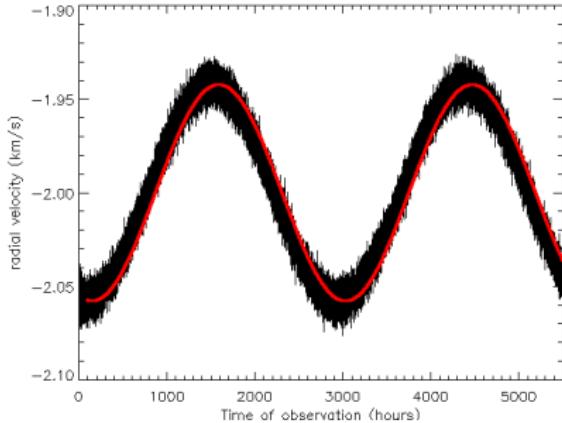
Trykk her!

Vi kan nå gjøre helt det samme for perioden P . Er du enig i at perioden (avstanden fra topp til topp) ihvertfall må være større enn avstanden mellom de røde linjene, og ihvertfall mindre enn avstanden mellom de blå linjene. Da har du en minste periode og en største periode som du kan lage f.eks. 20 mulige verdier mellom, akkurat som for t_0 . Og til slutt gjør du det samme for v_* . Da har du til slutt 20^3 mulige kombinasjoner.



Trykk her!

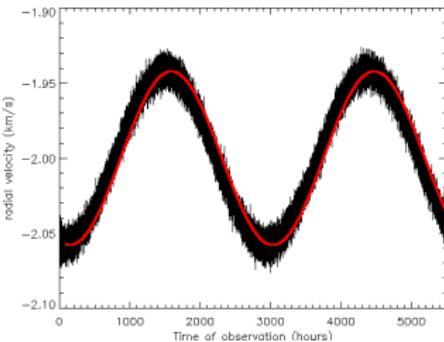
Når du så har funnet den av de 20^3 kombinasjonene som gir de beste estimatene av v_* , P og t_0 , så kan du tegne modellen $v_r^{\text{mod}}(t_i)$ og sammenlikne med de observerte dataene. Stemmer det godt overens? Det du spesielt skal tenke på her er om modellen din går midt inne i støyen, husk at støyen er Gaussisk slik at du forventer like mye utslag over som under kurven. Hva gjør du hvis du får et slikt resultat:



Passer denne modellkurven godt med dataene? Ligger modellkurven midt inne i støyen overalt? Hopper støyen like mye opp som ned overalt på kurven?

Trykk her når du har svaret!

NEI! Er du enig i at det **ikke** er tilfelle her? I noen deler av figuren så er støyen nesten hele tiden over eller under modellkurven, den opper ikke like mye over som under!



Her kan du være sikker på at noe er feil i koden! Men hvordan skal du finne feilen???? Her er det selvfølgelig to muligheter:

- Du har gjort noe feil når du valgte intervallene slik at den beste modellen ikke kommer med!
- Det er noe feil i koden

Aller først bør du dobbeltsjekke at du leste av riktig da du lagde prøve-intervallene for de forskjellige parameterene. Men hvis de er riktige, så må vi debugge... [Neste side](#)

- ① Normalt når kurven er så feil at du ser det tydelig slik som her, så er det ganske lett å finne en kurve som passer bedre på øyemål. Gjør så godt du kan for å lese av på figuren og finne en kombinasjon av v_* , P og t_0 som du mener passer godt og i hvertfall bedre enn det koden ga deg.
- ② Legg til et par linjer i koden som i tillegg til dine 20^3 kombinasjoner, også tester modellen som du fant ved øyemål. Print ut verdien for minste kvadrat (eller χ^2) og sammenlikn med verdien du får for modellen din. Gir modellen din et enda mindre kvadrat?
- ③ Nå skal du plotte din øyemål-modell over dataene, på samme måte som den røde linjen over. Passer din modell godt? Hvis nei, så ligger sannsynligvis feilen i måten du beregner $v_r^{\text{mod}}(t_i)$ ifra parameterene v_* , P og t_0 . Hvis du har gjort en god avlesning på øyemål, så **skal** du nå se en kurve som tilsynelatende ser ut til å passe bra. Hvis du har beregnet $v_r^{\text{mod}}(t_i)$ riktig.

- ④ Hvis øyemål-modellen din passer ganske bra, ga den en verdi for Δ eller χ^2 som er mindre enn minste kvadrat-verdien fra de 20^3 kombinasjonene? Hvis ikke, så er noe galt med måten du beregner Δ eller χ^2 på. Husk at disse størrelsene er forskjellen mellom data og modell. Hvis modellen din helt tydelig er bedre, så **må** den også ha en mindre forskjell, enig? Da må du sjekke hvordan du beregner denne forskjellen, hvorfor får du ikke mindre forskjell?
- ⑤ Hvis du får mindre forskjell for din modell, så er spørsmålet hvorfor ikke en minst like god modell ble funnet blandt de 20^3 kombinasjonene. Finn ut hvilken av de 20^3 kombinasjonene som er nærmest øyemålmodellen din. En enkel måte å gjøre det på er som følger: anta at du hadde 20 prøververdier for v_* , fra $v_*(0)$ som minste prøververdi og steglengde Δv_* slik at største prøververdi var $v_*(0) + 19 \times \Delta v_*$. Ta nå din øyemålverdi for v_* , del denne på Δv_* og rund av til nærmeste heltall. Da har du indeksen blant de 20 prøververdiene som passer best til din øyemålmodell.

- 6 Pass på: hvis din øyemålmodell er blant de aller første eller alle siste indeksene, altså at den er nær eller under $v_*(0)$, eller nær eller over $v_*(0) + 19 \times \Delta v_*$, så betyr det at du har valgt intervallet ditt galt slik at den beste løsningen ikke ligger omtrent midt inne i intervallet. Da er feilen rett og slett at du har lest av intervallet galt. Dette er en typisk feil!
- 7 Gjør det samme for P og t_0 . Hvis du fant at alle de 3 verdiene nå ligger omtrent midt i intervallet, så har du funnet den modellen bland de 20^3 modellene som er nærmest øyemålmodellen din. Nå skal du gjenta stegene over med denne nye testmodellen. Blir figuren pen med denne modellen også? (det bør den!). Blir Δ eller χ^2 mindre enn minste kvadrat fra koden også for denne modellen? (det bør den!). Hvis ja, hvorfor iallverden ble ikke denne modellen valgt som den beste av de 20^3 modellene, hvis denne modellen er identisk med en av de 20^3 modellene? Da ligger feilen enten i løkken din som gjør at ikke alle 20^3 modeller blir loopet gjennom, eller at det er noe galt i selve minimaliseringen.

Forrige side



Da er vi ferdige med del 1C. Vi har i dette kapittelet sett hvordan vi kan bruke stråling og spektrallinjer til å finne ut om en stjerne har planeter. I del 1D skal vi se nærmere på stråling generelt og spektrallinjer spesielt og den informasjonen som disse kan gi oss om gjerne objekter. I del 1C bør du nå kunne

- Vite hvordan man kan finne ut om en stjerne har en planet i bane rundt seg
- Forstå sammenhengen mellom en stjernes bane omkring massesenteret til en planet og Dopplereffekten som vi måler fra stjerna.
- Kunne bruke radialhastighetsgrafer til å anslå massen til planeten
- Kunne bruke flukskurve ved formørkelse til å anslå planetens radius og tetthet.
- Forstå hvordan støy påvirker data og hvordan man kan estimere parametere for en teoretisk modell fra data med støy.
- Forstå hvor minste kvadraters metode og χ^2 -metoden kommer fra og kunne bruke disse.

Neste side

[Forrige side](#)



Det anbefales nå at du sjekker [kortsvarsoppgavene](#) til del 1C for å kontrollere at du har forstått stoffet. Kan du svare på disse, blir det lettere å bruke kunnskapen din i oppgavene/prosjektet. Noen av disse kommer på eksamen.

Trykk nå gjerne på smilefjesene og si meningen din om dette interaktive forelesningsnotatet. Spesielt er jeg interessert i å vite hvor lang tid du brukte og da hva du brukte mest tid på! All ris og ros mottas med takk for å vite om alt arbeidet med å lage disse interaktive slidene er verd det, om jeg bør fortsette med det og om det er noe jeg bør endre.