Bayes reiknitækni

Helgi Tómasson

26. febrúar 2024

Dæmi í kafla 12.5, example 12.1

- Ef notast er við conjugate-prior verður bayesísk uppfærsla upplýsinga einföld formúla.
- Fyrir normal líkan er conjugate-prior fyrir meðaltalið, μ, normal dreifing.
- Fyrir normal líkan er conjugate-prior fyrir $v = 1/\sigma^2$ gammadreifing.
- Í dæminu í kafla 12.5 er Gibbs-samping notuð til að búa til Markov-keðju þar sem posterior-dreifingin er jafnvægisdreifing keðjunnar.
- Í praxís þarf heilmikla „diagnostics" til að greina hvort keðjan hafi náð jafnvægi.

- Parameter-vectorinn er:

$$\boldsymbol{\theta}' = [\beta \quad \phi_1 \quad \phi_2 \quad \sigma]$$

- Gibbs gengur út á að $\beta|\phi_1, \phi_2, \sigma$ sé normal, $\phi_1, \phi_2|\beta, \sigma$ sé normal (tvívítt) og rétt skalað $1/\sigma^2|\beta, \phi_1, \phi_2$ sé χ^2. Þ.e. σ^2 sé inverse-gamma.
- Hnitar vektorsins eru uppfærðar hverju skrefi og þannig fæst runa sem dreifist eins og posterior dreifingin.
- Hef forritað dæmi 12.1, bls. 625-627, skipanir í `ch_12_reg_time_error_3A.r` , útkoma:

```
                [,1]        [,2]       [,3]        [,4]        [,5]
lina1 -0.0001960954 0.793514453 0.18388808 -0.03652154 0.070631890
lina2  0.0016568417 0.007886523 0.02142856  0.02074809 0.001014506
```

- Athugið að ég er með fasta í líkani, bók ekki. Ég sýni meðaltal af σ, bók sýnir meðaltal af σ^2.
- Keðjan er í breytu sem heitir `gibbsurtak` .

Ný reiknitækni

- Í MCMC hermun þarf að sjúkdómsgreina útkomuna. Það er erfiðara að stilla Metropolis-Hastings rétt þannig að fólk hefur frekar viljað Gibbs.
- Á síðustu 10 árum hefur komið fram ný reiknitækni sem þykir hafa áhugaverða eiginleika, HMC (Hamiltonian-Monte-Carlo eða Hybrid-Monte-Carlo). Markov keðjurnar sem koma út úr slíkri hermun þykja hafa betri eiginleika (minni sjálffylgni ofl.) en eldir hermunaraðferðir.
- Reiknitækninni hefur verið pakkað í library sem er kallað stan. Stan er frekar „low-level" þannig að það er erfitt að forrita í því. Til ýmsir gluggar sem miða að því að gera það liprara í notkun.
- Leysti dæmi í kafla 12 með stan.

- Þarf stan skipanaskrá, `m4.stan`

```
data {
int<lower=1> n;
real x[n];
real y[n];
}
parameters {
real alpha;
real <lower=0> sigma;
real phi1;
real phi2;
}
transformed parameters {
real<lower=0> v;
v=10*0.05/sigma^2;
}
model {
alpha ~ normal(0,4);
phi1 ~ normal(0,0.25);
phi2 ~ normal(0,0.16);
v ~ gamma(20,0.5);
for (i in 3:n)
y[i] ~ normal(alpha*x[i]+phi1*(y[i-1]-alpha*x[i-1])+phi2*(y[i-2]-alpha*x[i-2]),sigma);
```

- Stúktúrinn er nokkrar skilgreiningablokkir.
- Síðan er kallað á þetta úr t.d. R.

- Nokkrar R-skipanir:

```
r1d=read.table("w-gs1yr.txt",header=T)
r1=r1d[,4]
r3d=read.table("w-gs3yr.txt",header=T)
r3=r3d[,4]
y=diff(r3)
x=diff(r1)
n=length(y)
library(rstan)
options(mc.cores = parallel::detectCores())
rstan_options(auto_write = TRUE)
model4=stan_model("m4.stan")
fit4=sampling(model4,list(n=n,y=y,x=x),iter=2000,chains=8)
```

Stan-útkoma

```
tafla121lina1 <- c(mean(unlist(extract(fit4)[1])),mean(unlist(extract(fit4)[3])),mean(unlist(extract(fit4)[4])),
mean(unlist(extract(fit4)[2])^2))

tafla121lina2 <-  c(var(unlist(extract(fit4)[1]))^0.5,var(unlist(extract(fit4)[3]))^0.5,var(unlist(extract(fit4)[4]))^0.5,
var(unlist(extract(fit4)[2])^2)^0.5)

tafla121 <- rbind(tafla121lina1,tafla121lina2)

tafla121
                     [,1]       [,2]        [,3]         [,4]
tafla121lina1 0.793578054 0.18163025 -0.03540972 0.0047411607
tafla121lina2 0.007576652 0.02036462  0.02033099 0.0001334352
```

Sama dreifing og í bók.

- Athugið að state-space reglurnar í kafla 11 nota reglu Bayes en eru ekki endilega Bayesísk aðferðafræði.
- Kalman-filter er hentug reikniverk við að fá tölulegt gildi á likelihood-föll.
- State-space aðferðafræði hentar við einvíð og margvíð vandamál. Allt sem hægt er að skrifa niður sem línulegar diffurjöfnur má skrifa á state-space formi (Athugið hvort þetta er rétt).
- Reikniverkið í kafla 8 um margvíð tímaraðalíkön mætti nálgast á state-space formi.
- Skoðun á ACF og PACF, auðveld í einni vídd, í mörgum víddum eru þetta runur af fylkjum.

Nokkur atriði úr kafla 8

- Hér er verið að útvíkka kafla 2.
- Vektor-ARMA eru miklu erfiðari en einvíð ARMA. Erfiðari í túlkun, margar framsetningar til á sama hreyfimynstri.

Tengsl hnita (breyta)

- Impulse-response fall, kerfinu varpað yfir í VMA(∞) og skoðað hvernig högg á einstakar jöfnur lekur í gegnum kerfið.
- Högg (impúlsar) geta verið korreleruð og því erfitt að álykta. Stundum reynt að umbreyta kerfi til að losna við fylgni. Erfitt að túlka umbreyttu stærðirnar.
- Stundum reynt að velja túlkanlega framsetningu, strúktúral VAR.
- Error-correction/co-integration og non-stationary kerfi.

ECM og trend-strúktúr

- Taka vel eftir umræðu um deterministíska þætti bls 434-435.
- Athugið:

$$y_t = \mu + \phi y_{t-1} + \varepsilon_t$$

  Ef $|\phi| <$ þá er þtetta stationary ferli með fast meðaltal.
- Ef $\phi = 1$, þá er $y_t = \mu t + \sum^t \varepsilon_t$, þ.e. random-walk með drift.
- Samanburður $|\phi| < 1$ og $\phi = 1$ er því samanburður á mjög ólíkum ferlum.
- Einhugmynd er að skilgreina:

$$y_t = \mu(\phi - 1) + \phi y_{t-1} + \varepsilon_t.$$

  Ef $\phi = 1$ hér þá er y_t random-walk, ef $|\phi| < 1$ þá er y_t stationary.
- Þetta dæmi ætti að skýra út samanburðinn á líkönunum á bls. 434-435, restricted-constant, restricted-trend, o.s.frv.

Unit-root í mörgum víddum

- Ef $\boldsymbol{y}_t$ er $I(1)$, margvíður random-walk. Þá má hugsa sér:

$$\Delta \boldsymbol{y}_t = \Pi \boldsymbol{y}_t + \varepsilon_t.$$

  Hvernig fylki er Π? Takið varíans í gegnum jöfnuna,

$$V(\Delta \boldsymbol{y}_t) = \Pi V(\boldsymbol{y}_t)\Pi' + \Sigma.$$

  Vinstri hliðin er endanleg. Hvað segir það um Π?
- Π hlýtur að vera einhvers konar núll-fylki því $V(\boldsymbol{y}_t)$ er ekki endanlegt.
- Þ.e. Π er ekki full rank. Segjum að $rank(\Pi) = r$, þar sem r er minna en vítt fylkisins. Skrifum $\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}'$, þar sem $\boldsymbol{\alpha}$ og $\boldsymbol{\beta}$ eru full rank-fylki.
- $V(\boldsymbol{\beta}'\boldsymbol{y}_t)$ hlýtur því að vera endanlegt.
- Ef svona Π er til eru hnitarnar í $\boldsymbol{y}_t$ co-integrated. Þ.e. $\boldsymbol{\beta}$ lýsir jafnvægissambandi og $\boldsymbol{\alpha}$ lýsir aðlögunarhraða að jafnvæginu.

- Setjum kerfið fram á error-correction formi (alltaf hægt):

$$\Delta \boldsymbol{y}_t = \boldsymbol{\mu} + \gamma t + \sum \Gamma_k \Delta \boldsymbol{y}_{t-k} + \boldsymbol{\alpha}\boldsymbol{\beta}' \boldsymbol{y}_{t-1} + \varepsilon_t,$$

- $\boldsymbol{\beta}$ myndar grunn fyrir co-integrating rúmið (r vítt) og $\boldsymbol{\alpha}$ lýsir að lögunarhraða að jafnvægi. Aðrir þættir taka á deterministískum þáttum og skammtímaþáttum.
- Hinn tölfræðilegi vandi lýtur að því að álykta um hvort kerfið sé $I(1)$ og að meta parametra út frá mælingum.

Dæmi

- Hugsum okkur tvívítt kerfi, íslensk og alþjóðlegt bensínverð. Þ.e. í mesta lagi 1 co-integrating samband. Höfum kerfi:

$$\begin{bmatrix} \Delta x_t \\ \Delta y_t \end{bmatrix} = \text{ýmsir þættir} + \begin{bmatrix} \alpha_x \\ \alpha_y \end{bmatrix} (y_{t-1} - \beta x_{t-1}) + \text{residual}$$

- Hvernig á að túlka þetta? er einhver breyta exogen?
- Hugsum okkur gengi gjaldmiðla, vexti og verðbólgu (5 breytur). Hugleiðum PPP og UAP. Hverjir gætu co-integrating vektorar verið? Hvað er co-integrating-rúmið vítt?
- Bayesísk nálgun erfið (held ég).