AALBORG UNIVERSITET

Tidsrækkeanalyse – Slides 7

Ege Rubak
(baseret på Slides af Esben Høg)

Institut for Matematiske Fag
Aalborg Universitet

Long memory tidsrækker

- En tidsrække har kort hukommelse (dvs. *short memory*) hvis

$$\sum |\gamma(h)| < \infty.$$

- En tidsrække hvor

$$\sum |\gamma(h)| = \infty,$$

  siges at have lang hukommelse (dvs. *long memory*).

Hvorfor er det relevant?

- Skriv gennemsnittet af $x_1, x_2, \dots, x_n$ som

$$\bar{x}_n = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

- Så er

$$\mathrm{V}(\bar{x}_n) = \frac{1}{n} \sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \gamma(h),$$

jvf. `Slides2.pdf` side 6.

Relevans (fortsat)

- Hvis $\sum |\gamma(h)| < \infty$, så er

$$\sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right) \gamma(h) \to \sum_{h=-\infty}^{\infty} \gamma(h) \text{ for } n \to \infty.$$

- Så

$$n\mathrm{V}(\bar{x}_n) \to \sum_{h=-\infty}^{\infty} \gamma(h),$$

$$\text{eller} \quad \mathrm{V}(\bar{x}_n) = \frac{1}{n}\left[\sum_{h=-\infty}^{\infty} \gamma(h)\right] + o\left(\frac{1}{n}\right).$$

Relevans (fortsat)

- Det vil sige, for en *short memory* tidsrække, så går $\mathrm{V}(\bar{x}_n)$ mod nul med den sædvanlige "hastighed" σ^2/n, når stikprøvestørrelsen stiger.
- Bemærk, at
$$\sum_{h=-\infty}^{\infty} \gamma(h) = f(0),$$
spektraltætheden $f(\omega)$ evalueret i $\omega = 0$.
- Så vi kan også skrive
$$\mathrm{V}(\bar{x}_n) = \frac{f(0)}{n} + o\left(\frac{1}{n}\right):$$
Den sædvanlige σ^2 er erstattet med $f(0)$.

Relevans (fortsat)

- **Men** hvis $\sum |\gamma(h)| = \infty$, så virker ovenstående ikke.
- I praksis, så viser mange tidsrækker, at $\mathrm{V}(\bar{x}_n)$ aftager mere langsomt. Plot $\log[\mathrm{V}(\bar{x}_n)]$ mod $\log(n)$, og kig efter en hældning på -1.

```
vartime = function(x, nmax = round(length(x) / 10)) {
v = rep(NA, nmax);
for (n in 1:nmax) {
y = filter(x, rep(1/n, n), sides = 1);
v[n] = var(y, na.rm = TRUE);
}
plot(log(1:nmax), log(v));
lmv = lm(log(v) ~ log(1:nmax));
abline(lmv);
title(paste(deparse(substitute(x)), "; nmax = ", nmax));
print(summary(lmv));
}
vartime(log(varve))
vartime(residuals(lm(log(varve) ~ time(log(varve)))))
```

Fraktionel Integration

- Hvordan kan man modellere sådanne serier, der har lang hukommelse?
- Fraktionel integreret hvid støj:

$$(1 - \mathrm{B})^d x_t = w_t, \quad 0 < d < 0.5.$$

- ACF viser sig at være (se senere)

$$\rho(h) = \frac{\Gamma(1-d)\Gamma(h+d)}{\Gamma(d)\Gamma(h+1-d)} \sim h^{2d-1},$$

  hvor $\Gamma(\cdot)$ er gammafunktionen
  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad \text{for } x \neq 0, -1, -2, \cdots.$
- Så for $0 < d < 0.5$,

$$\sum_{h=-\infty}^{\infty} |\rho(h)| = \infty.$$

Fraktionel Integration og lang hukommelse

- Nye begreber:
  - Fraktionel differensning.
  - Fraktionel integreret støj.
  - Fraktionelle integrerede processer.
- Vi skal se på hvad disse dækker over.

Fraktionel Integration og lang hukommelse (fortsat)

- Vi skal se på
  - Filteret bagved fraktionel differensning og integration.
  - De stokastiske egenskaber ved fraktionel integreret støj.
  - De stokastiske egenskaber ved fraktionel integrerede processer $\{x_t\}_{t\in\mathbb{Z}}$.

Fraktionel differensning

- Husk definitionen af en I(1) proces $\{y_t\}_{t\in\mathbb{Z}}$ ved hjælp af en I(0) proces $\{x_t\}_{t\in\mathbb{Z}}$ givet differensoperatoren $\nabla = 1 - \mathrm{B}$:

$$\nabla y_t = x_t, \quad t = 1, \dots, n.$$

- Ved inversion opnår man

$$y_t = y_0 + \sum_{j=0}^{t-1} x_{t-j}.$$

- Bemærk at "integrationen" stopper ved begyndelsen af stikprøven, fordi en proces som "$\sum_{j=0}^{\infty} x_{t-j}$" ikke eksisterer pga. ikke-stationaritet.
- Hvordan kan dette udvides til ∇^d med $d \notin \{0, 1\}$?

Fraktionel differensning

- I [ShuSt] afsnit 5.1 beskrives en lidt indforstået matematik vedrørende koefficienterne i rækkeudviklingen for de fraktionelle differenser, idet man betragter

$$\nabla^d = \sum_{j=0}^{\infty} \pi_j B^j = (1 - \mathrm{B})^d = 1 + \sum_{j=1}^{\infty} (-1)^j \binom{d}{j} \mathrm{B}^j. \qquad (1)$$

- I (1) indgår binomialkoefficienter $\binom{d}{j}$ hvor $j > d$ og herved indgår fakultetsfunktionen med negative og/eller ikke-heltallige argumenter. Disse egenskaber er dog veldefinerede, som det fremgår af det følgende.

Binomialrækker

- For d positiv og heltallig gælder en velkendt binomialformel:

$$(x+y)^d = \sum_{j=0}^{d} \binom{d}{j} x^j y^{d-j}. \tag{2}$$

- Denne formel kan skrives

$$(x+y)^d = \sum_{j=0}^{\infty} \binom{d}{j} x^j y^{d-j}, \tag{3}$$

hvis det underforstås, at $\binom{d}{j} = 0$ for d heltallig og $j > d$.

Binomialrækker (fortsat)

- En grundlæggende formel (som her tages for givet) fra uendelige rækker er en generalisering af (3), nemlig *den generelle binomialformel*, som viser rækkeudviklingen for potenser som ikke nødvendigvis er positive heltal:

$$(x+y)^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} x^j y^{\alpha-j}, \tag{4}$$

  hvor (4) skal forstås som (3), hvis α er et positivt heltal, hvorimod $\binom{\alpha}{j}$ er defineret på passende måde som vist nedenfor hvis α ikke er heltallig. Endvidere er det forudsat, at $x^2 < y^2$.
- Indsættes $x = -\mathrm{B}$, $y = 1$ og $\alpha = d$ fås netop (1).

Gammafunktionen og generaliserede binomialkoefficienter

- Gammafunktionen er formelt defineret som

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad \text{for } x \neq 0, -1, -2, \cdots$$

- Det ses let, at $\Gamma(1) = 1$, og det kan vises, at følgende rekursion gælder for alle x:

$$\Gamma(x+1) = x\Gamma(x). \tag{5}$$

- Specielt er $\Gamma(x+1) = x!$ for x heltallig og positiv. Man kan ikke finde et pænt lukket udtryk for gammafunktionen, men normalt er den tabelleret[1] i området $0 < x < 1$, hvorefter funktionsværdien i alle andre positive tal kan findes ved rekursionsformlen. For negative (ikke-heltallige) x kan funktionsværdierne fås ved omvendt brug af rekursionsformlen:

$$\Gamma(x) = \frac{1}{x}\Gamma(x+1). \tag{6}$$

[1] Eller der findes algoritmer, der kan beregne funktionsværdierne numerisk

Gammafunktionen og generaliserede binomialkoefficienter

- Gammafunktionen kan herefter bruges til at definere de *generaliserede binomialkoefficienter*.
- Som bekendt er for positive hele tal $d > j$ binomialkoefficienten givet ved

$$\binom{d}{j} = \frac{d!}{j!(d-j)!}.$$

- Omskrevet vha. gammafunktionen giver dette

$$\binom{d}{j} = \frac{\Gamma(d+1)}{\Gamma(j+1)\Gamma(d-j+1)}. \tag{7}$$

- Hvis nu j er positiv og heltallig og $|d| < 1$ så er $j - d > 0$ og rekursionsformlen (5) giver:

$$\begin{aligned}\Gamma(j-d) &= (j-d-1)\Gamma(j-d-1) \\ &\vdots \\ &= (j-d-1)(j-d-2)\cdots(-d)\Gamma(-d).\end{aligned} \tag{8}$$

- Endvidere er $d - j + 1 < 1$, hvorefter den omvendte rekursionsformel (6) giver

$$\begin{aligned}\Gamma(d-j+1) &= \frac{1}{d-j+1}\Gamma(d-j+2) \\ &\vdots \\ &= \frac{1}{d-j+1}\cdot\frac{1}{d-j+2}\cdots\frac{1}{d}\Gamma(d+1) \\ &= (-1)^j\frac{1}{j-d-1}\cdot\frac{1}{j-d-2}\cdots\frac{1}{-d}\Gamma(d+1) \\ &= (-1)^j\frac{\Gamma(d+1)\Gamma(-d)}{\Gamma(j-d)}.\end{aligned} \tag{9}$$

- Det næstsidste lighedstegn fås ved at gange samtlige j brøker med -1 i tællere og nævnere. Det sidste lighedstegn fås af (8).
- Hermed kan den generaliserede binomialkoefficient skrives (ved indsættelse af (9) i (7))

$$\binom{d}{j} = (-1)^j \frac{\Gamma(j-d)}{\Gamma(-d)j!}. \tag{10}$$

- Derfor fås altså (1) udtrykt som

$$\nabla^d = (1-\mathrm{B})^d = 1 + \sum_{j=1}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)j!}\mathrm{B}^j, \tag{11}$$

og dette er nøjagtig formel (5.2) i [ShuSt].

- Man ser, at

$$(-1)\binom{d}{1} = \frac{\Gamma(1-d)}{\Gamma(-d)} = \frac{(-d)\Gamma(-d)}{\Gamma(-d)} = -d.$$

- For $j \geq 2$ fås[2] følgende rekursion for koefficienterne

$$\begin{aligned}(-1)^{j+1}\binom{d}{j+1} &= (-1)^{2(j+1)}\frac{\Gamma(j+1-d)}{\Gamma(-d)(j+1)!} \\ &= \frac{j-d}{j+1}\cdot(-1)^{2j}\frac{\Gamma(j-d)}{\Gamma(-d)j!} \\ &= \frac{j-d}{j+1}\cdot(-1)^{j}\binom{d}{j}. \end{aligned} \tag{12}$$

[2]Bemærk, at $(-1)^{2(j+1)} = (-1)^{2j} = 1$

- Definerer vi $\pi_j(d) = (-1)^j \binom{d}{j}$, har vi altså at formel (12) svarer nøjagtigt til formel $(5.7)^3$ i [ShuSt]:

$$\pi_{j+1}(d) = \frac{j-d}{j+1}\pi_j(d).$$

- Herefter kan følgende tabel let laves:

j	$(-1)^j \binom{d}{j}$
1	$-d$
2	$(-1)^2 \binom{d}{2} = -1/2d(1-d)$
3	$(-1)^3 \binom{d}{3} = -1/2d(1-d)(2-d)/3 = -1/6d(1-d)(2-d)$
$\vdots$	$\vdots$

[3] Bemærk, at $\pi_j(d)$,erne er koefficienterne i AR(∞) repræsentationen af x_t.

Eksempel 1: Fraktionel integreret hvid støj

- Model

$$(1-\mathrm{B})^d x_t = w_t \qquad [\text{ARFIMA}(0,d,0)]. \tag{13}$$

For $|d| < \frac{1}{2}$ så er

$$(1 - d\mathrm{B} - \frac{1}{2}d(1-d)\mathrm{B}^2 - \frac{1}{6}d(1-d)(2-d)\mathrm{B}^3 - \cdots)x_t = w_t,$$

dvs.

$$x_t = dx_{t-1} + \frac{1}{2}d(1-d)x_{t-2} + \frac{1}{6}d(1-d)(2-d)x_{t-3} + \cdots + w_t,$$

eller skrevet som AR(∞) (idet $(1-\mathrm{B})^d = 1 + \sum_{j=1}^{\infty}(-1)^j\binom{d}{j}\mathrm{B}^j$)

$$\begin{aligned} x_t &= -\sum_{j=1}^{\infty}(-1)^j\binom{d}{j}x_{t-j} + w_t \\ &= -\sum_{j=1}^{\infty}\pi_j(d)x_{t-j} + w_t. \end{aligned}$$

Eksempel 1

- Alternativt kan man skrive processen som en MA(∞), jvf. [ShuSt] formel (5.4) og (5.5),

$$
\begin{aligned}
x_t = (1-\mathrm{B})^{-d} w_t &= \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}\, w_{t-j} \qquad (14) \\
&= \sum_{j=0}^{\infty} \psi_j(d)\, w_{t-j}.
\end{aligned}
$$

Autokorrelationer

- For processen givet i (13), hvor w_t er hvid støj og $|d| < \frac{1}{2}$, kan det vises, at x_t er stationær og invertibel, og at autokorrelationerne er givet ved

$$\rho(h) = \frac{\Gamma(1-d)\Gamma(h+d)}{\Gamma(d)\Gamma(h+1-d)}, \tag{15}$$

  jfv. [ShuSt] formel (5.6).

- Den nemmeste måde at indse formel (15) på er ved at skrive x_{t+h} vha. formel (14), og så beregne $\gamma(h) = \mathrm{E}\,(x_t\, x_{t+h})$.
- Endvidere er variansen

$$\mathrm{var}(x_t) = \frac{\Gamma(1-2d)}{(\Gamma(1-d))^2}.$$

Fraktionel integreret hvid støj (fortsat)

- Spektraltætheden for den fraktionelle integrerede hvide støj givet ved $(1-\mathrm{B})^d x_t = w_t$ kan vises at være

$$f(\omega) = \frac{\sigma_w^2}{[4\sin(\pi\omega)^2]^d},$$

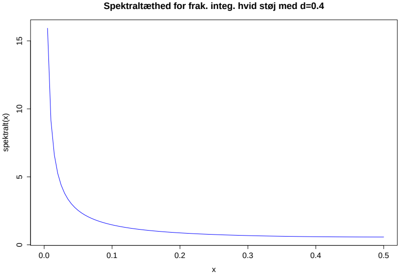
så for $d > 0$, $f(\omega) \to \infty$ når $\omega \to 0$.

Fraktionel integreret hvid støj (fortsat)

- Endvidere gælder $f(\omega) \sim |\omega|^{-2d}$ når $\omega \to 0$ så et plot af $\log[f(\omega)]$ mod $\log[|\omega|]$ giver et (præliminært) estimat af d.
- Hvis $d \geq 0.5$, så er $f(\omega)$ ikke integrabel, så tidsrækken er så ikke stationær.

Spektrum for fraktionel integreret hvid støj

```
sigma2 <- 1
d <- 0.4
spektralt <- function(omega)
sigma2/(4*sin(pi*omega)^2)^d
curve(spektralt, xlim = c(0, 0.5), col = "blue")
title(main="Spektraltæthed for frak. integ. hvid støj med d=0.4"
```



Eksempel 2

- Model

$$(1-\phi\mathrm{B})(1-\mathrm{B})^d x_t = w_t \qquad [\mathrm{ARFIMA}(1,d,0)]. \tag{16}$$

- Vi har

$$(1-\mathrm{B})^d = 1 + \sum_{j=1}^{\infty}(-1)^j \binom{d}{j} \mathrm{B}^j,$$

Eksempel 2 (fortsat)

- Derfor får vi
  $(1-\phi B)(1-B)^d =$

$$= 1 + \sum_{j=1}^{\infty} (-1)^j \binom{d}{j} B^j - \phi B - \sum_{j=1}^{\infty} \phi(-1)^j \binom{d}{j} B^{j+1}$$

$$= 1 + \sum_{j=1}^{\infty} (-1)^j \binom{d}{j} B^j - \phi B - \sum_{j=2}^{\infty} \phi(-1)^{j-1} \binom{d}{j-1} B^j$$

$$= 1 + \sum_{j=1}^{\infty} (-1)^j \binom{d}{j} B^j - \phi B - \sum_{j=2}^{\infty} \phi(-1)^{j-1} \frac{j}{d-j+2} \binom{d}{j} B^j$$

$$= 1 - (\phi + d)B - \sum_{j=2}^{\infty} \left[(-1)^j - \phi(-1)^{j-1} \frac{j}{d-j+2} \right] \binom{d}{j} B^j$$

$$= 1 - (\phi + d)B - (1 + \frac{2\phi}{d}) \binom{d}{2} B^2 - \cdots,$$

Eksempel 2 (fortsat)

- Alt i alt

$$\left(1-(\phi+d)\mathrm{B}-(1+\frac{2\phi}{d})\binom{d}{2}\mathrm{B}^2-\cdots\right)x_t=w_t.$$

- Det vil sige

$$x_t=(\phi+d)x_{t-1}+(1+\frac{2\phi}{d})\binom{d}{2}x_{t-2}+\cdots+w_t,$$

som er AR(∞) repræsentationen af x_t.

- Også her kan autokorrelationer og varians beregnes, når $|d|<\frac{1}{2}$.

ARFIMA model

- I nogle long-memory tidsrækker, så vil autokorrelationer ved små lags ikke matche dem der svarer til fraktionel integreret hvid støj. (Fx ligesom i Eksempel 2 ovenfor)
- Man kan så tilføje ARMA komponenter for at få tilpasset sådanne tidsrækker: ARFIMA(p, d, q) med d fraktionel, $-0.5 < d < 0.5$.
- Vi har til rådighed R funktionen `fracdiff()`.

Fraktionel Integreret tidsrække: ARFIMA

ARFIMA(p, d, q)

For ikke-negative heltal p, q, $-0.5 < d < 0.5$, siger vi at tidsrækken $\{x_t\}_{t \in \mathbb{Z}}$ er en ARFIMA(p, d, q) proces, hvis $y_t = \nabla^d x_t = (1 - \mathrm{B})^d x_t$ er ARMA(p, q).