AALBORG UNIVERSITET

Tidsrækkeanalyse – Slides 6

Ege Rubak
(baseret på Slides af Esben Høg)

Institut for Matematiske Fag
Aalborg Universitet

Spektralanalyse

- Idé: Dekomponér en stationær tidsrække $\{x_t\}_{t \in \mathbb{Z}}$ i en sum/kombination af sinussvingninger med tilfældige (og ukorrelerede) koefficienter.
- Ligeom i Fourieranalyse, hvor man dekomponerer deterministiske funktioner i kombinationer af sinussvingninger.
- Dette kaldes spektralanalyse eller analyse i frekvensdomænet, i modsætning til tidsdomænetilgangen, som vi har kigget på hidtil.
- Frekvensdomæne-tilgangen betragter så at sige regressioner på cosinus- og sinusfunktioner. Hvorimod tidsdomæne-tilgangen handler om regression på tidligere værdier (*lags*) af tidsrækken.

En periodisk tidsrække

Betragt

$$\begin{aligned} x_t &= a\sin(2\pi\omega t) + b\cos(2\pi\omega t) \\ &= c\sin(2\pi\omega t + \phi), \end{aligned}$$

hvor a og b er ukorrelerede med middelværdi 0 og varians 1, og $c^2 = a^2 + b^2$, $\tan\phi = b/a$. Så er

$$\begin{aligned} \mu_t &= \mathrm{E}\left[x_t\right] = 0, \\ \gamma(t, t+h) &= \cos(2\pi\omega h). \end{aligned}$$

Så vi kan konkludere, at $\{x_t\}_{t\in\mathbb{Z}}$ er stationær.

En sidebemærkning: Nogle trigonometriske identiteter

$$
\begin{aligned}
\tan\theta &= \frac{\sin\theta}{\cos\theta},\\
\sin^2\theta + \cos^2\theta &= 1,\\
\sin(a+b) &= \sin a\cos b + \cos a\sin b,\\
\cos(a+b) &= \cos a\cos b - \sin a\sin b,\\
e^{i\theta} &\equiv \cos\theta + i\sin\theta, \qquad \text{Eulers formel.}
\end{aligned}
$$

En periodisk tidsrække

- For $x_t = a\sin(2\pi\omega t) + b\cos(2\pi\omega t)$ med ukorrelerede a og b (middelværdier 0 og varianser σ^2), så er $\gamma(h) = \sigma^2\cos(2\pi\omega h)$.
- Autokovarianserne af summen af to ukorrelerede tidsrækker er summen af deres autokovarianser. Altså, er autokovariansen af en sum af tilfældige sinussvingninger lig med summen af sinussvingningerne med de tilsvarende frekvenser:

$$x_t = \sum_{j=1}^{k} \left(a_j \sin(2\pi\omega_j t) + b_j \cos(2\pi\omega_j t)\right)$$

$$\gamma(h) = \sum_{j=1}^{k} \sigma_j^2 \cos(2\pi\omega_j h),$$

hvor a_j og b_j er ukorrelerede med middelværdier 0, og $\mathrm{V}(a_j) = \mathrm{V}(b_j) = \sigma_j^2$.

En periodisk tidsrække (fortsat)

Altså, for en tidsrække på denne form:

- $x_t = \sum_{j=1}^{k} \left(a_j \sin(2\pi\omega_j t) + b_j \cos(2\pi\omega_j t)\right)$ har vi, at $\gamma(h) = \sum_{j=1}^{k} \sigma_j^2 \cos(2\pi\omega_j h)$.
- Altså kan vi repræsentere $\gamma(h)$ som en Fourierrække. Koefficienterne er varianserne af sinussvingnings-komponenterne.

Den spektrale fordelingsfunktion

For $x_t = a\sin(2\pi\omega_0 t) + b\cos(2\pi\omega_0 t)$ for en fast frekvens ω_0, har vi, jvf. slide 5, at $\gamma(h) = \sigma^2\cos(2\pi\omega_0 h)$, og vi kan skrive

$$\gamma(h) = \int_{-1/2}^{1/2} \mathrm{e}^{2\pi i\omega h}\mathrm{d}F(\omega),$$

hvor F er den diskrete fordeling

$$F(\omega) = \begin{cases} 0 & \text{hvis } \omega < -\omega_0, \\ \frac{\sigma^2}{2} & \text{hvis } -\omega_0 \leq \omega < \omega_0, \\ \sigma^2 & \text{ellers.} \end{cases}$$

Funktionen F opfører sig som en fordelingsfunktion for en diskret stokastisk variabel, bortset fra at $F(\infty) = \sigma^2 = \mathrm{V}(x_t)$ i stedet for 1. Man kan sige, at $F(\omega)$ er en fordelingsfunktion, ikke af sandsynligheder, men af varianser knyttet til frekvensen ω_0 i en analyse af variansen. Derfor kalder man $F(\omega)$ for den spektrale fordelingsfunktion.

Den spektrale fordelingsfunktion

Den spektrale fordelingsfunktion

For enhver staionær $\{x_t\}_{t\in\mathbb{Z}}$ med autokovarians γ, så kan vi skrive

$$\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i\omega h} dF(\omega),$$

hvor F er *den spektrale fordelingsfunktion* for $\{x_t\}_{t\in\mathbb{Z}}$.

Hvis γ er absolut summabel[1], så er F kontinuert: $dF(\omega) = f(\omega)d\omega$.
Hvis γ er en sum af sinusoider, så er F diskret.

[1] Det vil sige $\sum_{-\infty}^{\infty} |\gamma(h)| < \infty$.

Den spektrale fordelingsfunktion (fortsat)

Hvis $\{x_t\}_{t\in\mathbb{Z}}$ er den deterministiske proces $x_t = \sum_{j=1}^{k} (a_j \sin(2\pi\omega_j t) + b_j \cos(2\pi\omega_j t))$, så er den spektrale fordelingsfunktion af $\{x_t\}_{t\in\mathbb{Z}}$

$$F(\omega) = \sum_{j=1}^{k} \sigma_j^2 F_j(\omega),$$

hvor

$$F_j(\omega) = \begin{cases} 0 & \text{hvis } \omega < -\omega_j, \\ \frac{1}{2} & \text{hvis } -\omega_j \le \omega < \omega_j \\ 1 & \text{ellers.} \end{cases}$$

Spektraltætheden

Spektraltætheden

Hvis en tidsrække $\{x_t\}_{t\in\mathbb{Z}}$ har en autokovarians γ der tilfredsstiller $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$, så definerer vi spektraltætheden for $\{x_t\}_{t\in\mathbb{Z}}$ som

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h},$$

for $-\infty < \omega < \infty$.

Spektraltætheden: Nogle fakta

1. Vi har at $\sum_{h=-\infty}^{\infty} |\gamma(h)| e^{-2\pi i \omega h} < \infty$.
   Dette er fordi
   $|e^{i\theta}| = |\cos\theta + i\sin\theta| = (\cos^2\theta + \sin^2\theta)^{1/2} = 1$, og fordi γ er såkaldt absolut summabel.
2. f er periodisk med periode 1.
   Dette gælder fordi $e^{-2\pi i \omega h}$ er en periodisk funktion af ω med periode 1. Dvs. vi kan restringere domænet af f til $-1/2 \leq \omega \leq 1/2$ (som Shumway & Stoffer gør det).

Spektraltætheden: Nogle fakta

3. f er en lige funktion (dvs. $f(\omega) = f(-\omega)$). For at indse dette, skriv

$$
\begin{aligned}
f(\omega) &= \sum_{h=-\infty}^{-1} \gamma(h) e^{-2\pi i \omega h} + \gamma(0) + \sum_{h=1}^{\infty} \gamma(h) e^{-2\pi i \omega h}, \\
f(-\omega) &= \sum_{h=-\infty}^{-1} \gamma(h) e^{-2\pi i \omega (-h)} + \gamma(0) + \sum_{h=1}^{\infty} \gamma(h) e^{-2\pi i \omega (-h)} \\
&= \sum_{h=1}^{\infty} \gamma(-h) e^{-2\pi i \omega h} + \gamma(0) + \sum_{-\infty}^{-1} \gamma(-h) e^{-2\pi i \omega h} \\
&= f(\omega).
\end{aligned}
$$

4. $f(\omega) \geq 0$.

Spektraltætheden: Nogle fakta

5. $\gamma(h) = \int_{-1/2}^{1/2} e^{2\pi i \omega h} f(\omega) d\omega$. Hvorfor ?

$$\begin{aligned}
\int_{-1/2}^{1/2} e^{2\pi i \omega h} f(\omega) d\omega &= \int_{-1/2}^{1/2} \sum_{j=-\infty}^{\infty} \gamma(j) e^{-2\pi i \omega (j-h)} d\omega \\
&= \sum_{j=-\infty}^{\infty} \gamma(j) \int_{-1/2}^{1/2} e^{-2\pi i \omega (j-h)} d\omega \\
&= \gamma(h) + \sum_{j \neq h} \frac{\gamma(j)}{2\pi i (j-h)} \left(e^{\pi i (j-h)} - e^{-\pi i (j-h)} \right) \\
&= \gamma(h) + \sum_{j \neq h} \frac{\gamma(j) \sin(\pi(j-h))}{\pi (j-h)} = \gamma(h).
\end{aligned}$$

Sidste lighedstegn, fordi $\sin(\pi(j-h)) = 0$, da $j-h$ er hel.

Autokovarians og spektrum

- Bemærk, at $\gamma(0) = \mathrm{V}(x_t) = \int_{-1/2}^{1/2} f(\omega)\mathrm{d}\omega$, som udtrykker at total varians er integreret spektraltæthed over alle frekvenser.
- Autokovariansfunktionen $\gamma(h)$ og spektraltætheden indeholder samme information. Autokovariansen indeholder information i form af *lags*, hvorimod spektraltætheden indeholder den samme information i form af cyklusser.
- Nogle probelemer løses bedst ved at undersøge lagged information, og derfor arbejder man i tidsdomænet. Andre problemer løses måske bedst ved at betragte periodisk information, og der arbejder man så i frekvensdomænet (spektraldomænet).
- Bemærk at autokovariansfunktionen $\gamma(h)$ og spektraltætheden f er Fouriertransform-par (hinandens inverse).

Eksempel: Hvid støj

- For hvid støj $\{w_t\}_{t \in \mathbb{Z}}$ har vi set at $\gamma(0) = \sigma_w^2$ og $\gamma(h) = 0$ for $h \neq 0$.
- Dvs.

$$f(\omega) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \omega h} \tag{1}$$
$$= \gamma(0) = \sigma_w^2. \tag{2}$$

- Det vil sige, at spektraltætheden er konstant over alle frekvenser: Hver frekvens i spektret bidrager lige meget til variansen. Det er (som også tidligere nævnt) oprindelsen til betegnelsen *hvid støj*: Det er ligesom med hvidt lys, som er et ligeligt mix af alle frekvenser (farver) i det synlige spektrum.

Eksempel: AR(1)

For AR(1)-processen $x_t = \phi_1 x_{t-1} + w_t$, har vi set, at

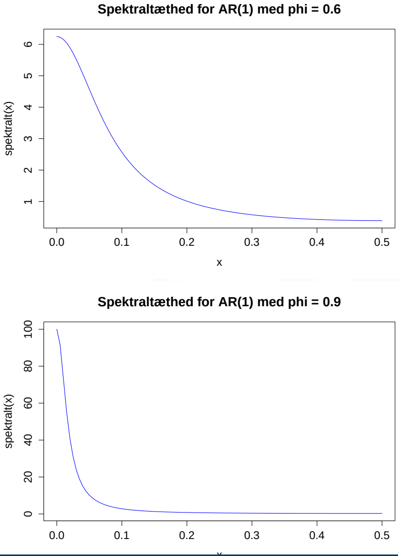
$$\gamma(h) = \sigma_w^2 \phi_1^{|h|} / (1 - \phi_1^2).$$

Det kan vises (gøres i en opgave) at

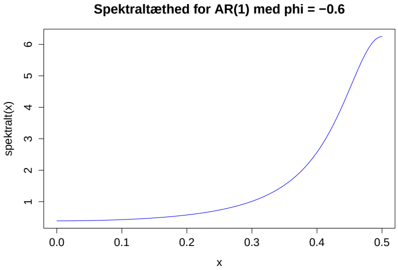
$$f(\omega) = \frac{\sigma_w^2}{1 - 2\phi_1 \cos(2\pi\omega) + \phi_1^2}.$$

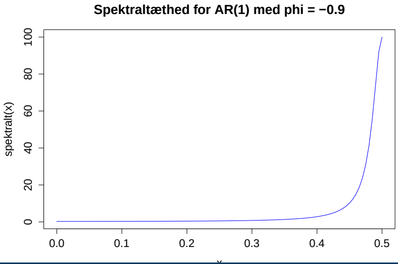
- Hvis $\phi_1 > 0$ (positiv autokorrelation), så er spektret domineret af lave frekvenser – dvs. forholdsvis glat i tidsdomænet.
- Hvis $\phi_1 < 0$ (negativ autokorrelation), så er spektret domineret af høje frekvenser – dvs. forholdsvis groft (ikke glat) i tidsdomænet.

Eksempel: AR(1)



Eksempel: AR(1)

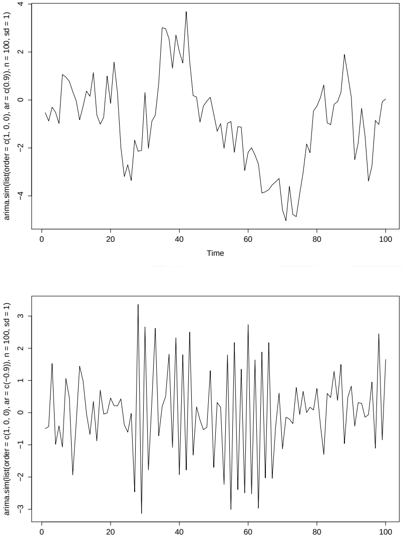




Simulation af AR(1) i tidsdomænet

```
plot.ts(arima.sim(list(ma = c(0.9)), n = 100))
plot.ts(arima.sim(list(ma = c(-0.9)),n = 100))
```

Simulation af AR(1) i tidsdomænet



Eksempel: MA(1)

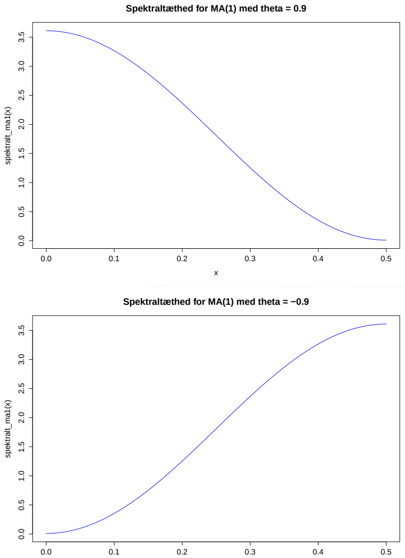
$$x_t = w_t + \theta_1 w_{t-1}.$$

$$\gamma(h) = \begin{cases} \sigma_w^2(1+\theta_1^2) & \text{hvis } h = 0, \\ \sigma_w^2\theta_1 & \text{hvis } |h| = 1, \\ 0 & \text{ellers.} \end{cases}$$

$$f(\omega) = \sigma_w^2 \left(1 + \theta_1^2 + 2\theta_1 \cos(2\pi\omega)\right).$$

- Hvis $\theta_1 > 0$ (positiv autokorrelation), så er spektret også her domineret af lave frekvenser – dvs. forholdsvis glat i tidsdomænet.
- Hvis $\theta_1 < 0$ (negativ autokorrelation), så er spektret også her domineret af høje frekvenser – dvs. forholdsvis groft (ikke glat) i tidsdomænet.

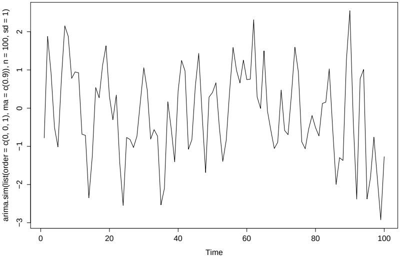
Eksempel: MA(1)

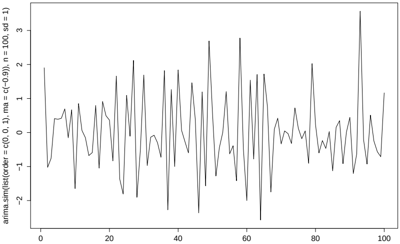


Simulation i tidsdomænet

```
plot(arima.sim(list(ma = 0.9), n = 100))
plot(arima.sim(list(ma = -0.9), n = 100))
```

Simulation af MA(1) i tidsdomænet





Den autokovarians-frembringende funktion og spektret

- Antag $\{x_t\}_{t\in\mathbb{Z}}$ er en kausal lineær proces, så den altså kan skrives som
$$x_t = \sum_{i=0}^{\infty} \psi_i w_{t-i} = \psi(\mathrm{B}) w_t.$$
- Vi har tidligere set at autokovariansfunktionen er
$$\gamma(h) = \mathrm{Cov}(x_t, x_{t+h}) = \sigma_w^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+h}.$$
- Definér nu den såkaldte autokovarians frembringende funktion som
$$\gamma(\mathrm{B}) = \sum_{h=-\infty}^{\infty} \gamma(h) \mathrm{B}^h.$$

Den autokovarians-frembringende funktion og spektret

Eftersom

$$\gamma(\mathrm{B}) = \sum_{h=-\infty}^{\infty} \gamma(h)\mathrm{B}^h,$$

så er

$$\begin{aligned}\gamma(\mathrm{B}) &= \sigma_w^2 \sum_{h=-\infty}^{\infty} \sum_{i=0}^{\infty} \psi_i \psi_{i+h} \mathrm{B}^h \\ &= \sigma_w^2 \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \mathrm{B}^{j-i} \\ &= \sigma_w^2 \sum_{i=0}^{\infty} \psi_i \mathrm{B}^{-i} \sum_{j=0}^{\infty} \psi_j \mathrm{B}^j = \sigma_w^2 \psi(\mathrm{B}^{-1})\psi(\mathrm{B}).\end{aligned}$$

Den autokovarians-frembringende funktion og spektret

Bemærk så at igen eftersom

$$\gamma(\mathsf{B}) = \sum_{h=-\infty}^{\infty} \gamma(h)\mathsf{B}^h,$$

så kan spektraltætheden (se slide 10) skrives

$$\begin{aligned} f(\omega) &= \sum_{h=-\infty}^{\infty} \gamma(h)\mathrm{e}^{-2\pi i\omega h} \\ &= \gamma(\mathrm{e}^{-2\pi i\omega}) \\ &= \sigma_w^2 \psi(\mathrm{e}^{-2\pi i\omega})\psi(\mathrm{e}^{2\pi i\omega}) \\ &= \sigma_w^2 |\psi(\mathrm{e}^{2\pi i\omega})|^2. \end{aligned}$$

Spektret for en MA(q)

For eksempel, for en MA(q) har vi, $\psi(\mathrm{B}) = \theta(\mathrm{B})$, så

$$\begin{aligned} f(\omega) &= \sigma_w^2 \theta(\mathrm{e}^{-2\pi i \omega}) \theta(\mathrm{e}^{2\pi i \omega}) \\ &= \sigma_w^2 |\theta(\mathrm{e}^{-2\pi i \omega})|^2. \end{aligned}$$

Specielt for en MA(1),

$$f(\omega) = \sigma_w^2 (1 + 2\theta_1 \cos(2\pi\omega) + \theta_1^2).$$

Spektret for en AR(p)

For en AR(p) har vi, $\psi(\mathrm{B}) = 1/\phi(\mathrm{B})$, så

$$f(\omega) = \frac{\sigma_w^2}{\phi(\mathrm{e}^{-2\pi i\omega})\phi(\mathrm{e}^{2\pi i\omega})} = \frac{\sigma_w^2}{|\phi(\mathrm{e}^{-2\pi i\omega})|^2}.$$

Specielt for en AR(1),

$$f(\omega) = \frac{\sigma_w^2}{(1 - 2\phi_1 \cos(2\pi\omega) + \phi_1^2)}.$$

Spektraltætheden for en kausal lineær proces

Hvis $\{x_t\}_{t\in\mathbb{Z}}$ er en kausal lineær proces, dvs. så kan den jo skrives $x_t = \sum_0^\infty \psi_i w_{t-i} = \psi(\mathrm{B})w_t$. Så har vi

$$f(\omega) = \sigma_w^2 |\psi(\mathrm{e}^{-2\pi i\omega})|^2.$$

Det vil sige, at spektraltætheden $f(\omega)$ for en lineær proces måler modulus af MA(∞)-polynomiet ψ i punktet $\mathrm{e}^{-2\pi i\omega}$ på enhedscirklen.

Spektraltætheden for en lineær proces (fortsat)

For en ARMA(p, q), $\psi(\mathrm{B}) = \theta(\mathrm{B})/\phi(\mathrm{B})$, så

$$
\begin{aligned}
f(\omega) &= \sigma_w^2 \frac{\theta(\mathrm{e}^{-2\pi i\omega})\theta(\mathrm{e}^{2\pi i\omega})}{\phi(\mathrm{e}^{-2\pi i\omega})\phi(\mathrm{e}^{2\pi i\omega})} \\
&= \sigma_w^2 \left| \frac{\theta(\mathrm{e}^{-2\pi i\omega})}{\phi(\mathrm{e}^{-2\pi i\omega})} \right|^2 .
\end{aligned}
$$

Denne er kendt som et *rationalt spektrum*.

Rationelle spektre

Betragt faktoriseringerne af $\theta(z)$ og $\phi(z)$:

$$\theta(z) = \theta_q(z - z_1)(z - z_2)\cdots(z - z_q),$$
$$\phi(z) = \phi_p(z - p_1)(z - p_2)\cdots(z - p_p).$$

Altså $z_1, \ldots, z_q$ er rødderne (nulpunkterne) for $\theta(\cdot)$, og $p_1, \ldots, p_p$ er rødderne (nulpunkterne) for $\phi(\cdot)$.

Med hensyn til spektret $f(\omega) = \sigma_w^2 \left|\frac{\theta(e^{-2\pi i\omega})}{\phi(e^{-2\pi i\omega})}\right|^2$ for ARMA(p, q), så kalder man $z_1, \ldots, z_q$ for *nulpunkterne* for $f(\omega)$, og $p_1, \ldots, p_p$ kaldes for *polerne* for $f(\omega)$,

$$\begin{aligned} f(\omega) &= \sigma_w^2 \left|\frac{\theta_q \prod_{j=1}^q (e^{-2\pi i\omega} - z_j)}{\phi_p \prod_{j=1}^p (e^{-2\pi i\omega} - p_j)}\right|^2 \\ &= \sigma_w^2 \frac{\theta_q^2 \prod_{j=1}^q |e^{-2\pi i\omega} - z_j|^2}{\phi_p^2 \prod_{j=1}^p |e^{-2\pi i\omega} - p_j|^2}. \end{aligned}$$

Rationelle spektre

- Vi så altså, at

$$f(\omega) = \sigma_w^2 \frac{\theta_q^2 \prod_{j=1}^{q} |e^{-2\pi i\omega} - z_j|^2}{\phi_p^2 \prod_{j=1}^{p} |e^{-2\pi i\omega} - p_j|^2}.$$

- Som ω varierer fra 0 til 1/2, så bevæger $e^{-2\pi i\omega}$ sig med uret rundt på enhedscirklen fra 1 til $e^{-\pi i} = -1$.
- Og værdien af $f(\omega)$ stiger jo mere $e^{-2\pi i\omega}$ bevæger sig tættere på (længere væk fra) polerne p_j (nulpunkterne z_j).

Eksempel: ARMA

- Genkald AR(1): $\phi(z) = 1 - \phi_1 z$. Polen er her i punktet $p_1 = 1/\phi_1$.
  - Hvis $0 < \phi_1 < 1$, så er polen større end 1, så spektraltætheden aftager som ω bevæger sig væk fra 0.
  - Hvis $-1 < \phi_1 < 0$, så er polen mindre end -1, så spektraltætheden er ved sit maksimum, når $\omega = 0.5$.
- Genkald MA(1): $\theta(z) = 1 + \theta_1 z$. Nulpunktet er her i punktet $z_1 = -1/\theta_1$.
  - Hvis $0 < \theta_1 < 1$, så er nulpunktet mindre end -1, så spektraltætheden aftager som ω bevæger sig mod -1.
  - Hvis $-1 < \theta_1 < 0$, så er nulpunktet større end 1, så spektraltætheden er ved sit minimum, når $\omega = 0$.

Eksempel: AR(2)

Betragt $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + w_t$. Eksempel 4.7 i Shumway & Stoffer betragter denne model med $\phi_1 = 1$, $\phi_2 = -0.9$ og $\sigma_w^2 = 1$. I dette tilfælde er polerne tilnærmelsesvis

$$p_1, p_2 \approx 0.5555 \pm i0.8958 \approx 1.054e^{\pm i1.01567} \approx 1.054e^{\pm 2\pi i 0.16165}.$$

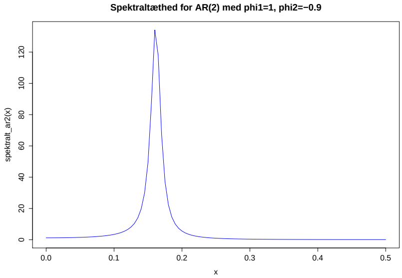
Det vil sige, vi har

$$f(\omega) = \frac{\sigma_w^2}{\phi_2^2 |e^{-2\pi i\omega} - p_1|^2 |e^{-2\pi i\omega} - p_2|^2},$$

og denne er meget *peaked* (stejl), når $e^{-2\pi i\omega}$ er tæt på $1.054e^{-2\pi i 0.16165}$.

Eksempel: AR(2)

```
sigma2 <- 1
phi1 <- 1
phi2 <- -0.9
arma.spec(ar=c(phi1, phi2), var.noise = sigma2, log = "no",
main = "AR(2) spektraltæthed (phi1=1, phi2=-0.9)")
```



Eksempel: Sæson ARMA

Betragt $x_t = \Phi_1 x_{t-12} + w_t$.

$$\psi(\mathrm{B}) = \frac{1}{1 - \Phi_1 \mathrm{B}^{12}},$$

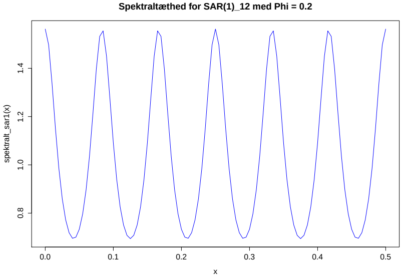
$$f(\omega) = \frac{\sigma_w^2}{(1 - \Phi_1 e^{-2\pi i 12\omega})(1 - \Phi_1 e^{2\pi i 12\omega})}$$

$$= \frac{\sigma_w^2}{1 - 2\Phi_1 \cos(24\pi\omega) + \Phi_1^2}.$$

Bemærk, at $f(\omega)$ er periodisk med periode $1/12$.

Eksempel: Sæson ARMA

```
sigma2 <- 1
Phi1 <- 0.2
arma.spec(ar=c(rep(0,11), Phi1), var.noise = sigma2, log = "no",
main = "SAR(1)_12 spektraltæthed (Phi=0.2)")
```



Eksempel: Sæson ARMA

En anden måde at se det på:

$$1 - \Phi_1 z^{12} = 0 \quad \Leftrightarrow \quad z = r e^{i\theta},$$
$$\text{med } r = |\Phi_1|^{-1/12}, \quad e^{i12\theta} = e^{-i \arg(\Phi_1)}.$$

For $\Phi_1 > 0$, så er de tolv poler givet ved $|\Phi_1|^{-1/12} e^{ik\pi/6}$ for $k = 0 \pm 1, \ldots, \pm 5, 6$.
Så spektraltætheden har *peaks*, når $e^{-2\pi i\omega}$ passerer igennem de tolv punkter

$$\{1, e^{\pm i\pi/6}, e^{\pm i\pi/3}, e^{\pm i\pi/2}, e^{\pm i2\pi/3}, e^{\pm i5\pi/6}, -1\}.$$