Solutions to exercises in

Stationary stochastic processes for scientists and engineers

by Lindgren, Rootzén and Sandsten
Chapman & Hall/CRC, 2013

Georg Lindgren, Johan Sandberg,
Maria Sandsten

2018

LUND
UNIVERSITY

Faculty of Engineering
Centre for Mathematical Sciences
Mathematical Statistics

CENTRUM SCIENTIARUM MATHEMATICARUM

Solutions to exercises in

Stationary stochastic processes for scientists and engineers

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118
SE-221 00 Lund, Sweden

http://www.maths.lu.se

Contents

Preface

This booklet contains hints and solutions to exercises in *Stationary stochastic processes for scientists and engineers* by Georg Lindgren, Holger Rootzén, and Maria Sandsten, Chapman & Hall/CRC, 2013.

The solutions have been adapted from course material used at Lund University on first courses in stationary processes for students in engineering programs as well as in mathematics, statistics, and science programs.

The web page for the course during the fall semester 2013 gives an example of a schedule for a seven week period:

`http://www.maths.lu.se/matstat/kurser/fms045mas210/`

Note that the chapter references in the material from the Lund University course do not exactly agree with those in the printed volume.

Chapter 2

Stationary processes

2:1. (a) 1, (b) $a+b$, (c) 13, (d) a^2+b^2, (e) a^2+b^2, (f) 1.

2:2. $\rho[X_1, X_2] = -\frac{4.5}{\sqrt{50}}$.

2:3. The mean value function: $m_X(t) \triangleq \mathsf{E}[X(t)] = \mathsf{E}[1.2e_t + 0.9e_{t-1}] = 1.2m + 0.9m = 2.1m$. The covariance function:

$$\begin{aligned} r_X(s,t) &\triangleq \mathsf{C}[X(s), X(t)] = \mathsf{C}[1.2e_s + 0.9e_{s-1}, 1.2e_t + 0.9e_{t-1}] \\ &= \begin{cases} (1.2^2 + 0.9^2)\sigma^2 = 2.25\sigma^2, & s = t, \\ 1.2 \cdot 0.9\sigma^2 = 1.08\sigma^2, & s = t \pm 1, \\ 0, & |s-t| \geq 2. \end{cases} \end{aligned}$$

A process $X(t)$ is weakly stationary if the mean value function, $m_X(t)$, does not depend on t *and* the covariance function, $r_X(t,s)$, only depends on $|t-s|$. Here the mean value function does not depend on t and the covariance function only depends on $|s-t|$ (check this by proving that $r_X(s+c,t+c) = r_X(s,t)$ for any c.). Thus, the process is weakly stationary.

2:4. $r_Y(1,2) = 0.5\sigma^2 \neq r_Y(2,3) = 0.625\sigma^2 \Rightarrow$ non-stationary. With $c = \sqrt{4/3}$ one gets a stationary sequence.

2:5. Hint: Stockholm has a temperate climate with four distinct seasons.

2:6. The expected value is given by

$$\mathsf{E}[X(t)] = \mathsf{E}[Y]\sin(t) + \mathsf{E}[Z] + 5 = 5.$$

The variance is

$$\begin{aligned} \mathsf{V}[X(t)] &= \mathsf{V}[Y\sin(t) + Z + 5] = \mathsf{C}[Y\sin(t), Y\sin(t)] + \mathsf{C}[Z,Z] \\ &= \sigma_y^2 \sin^2(t) + \sigma_z^2. \end{aligned}$$

The expected value is constant but the variance depends on t; thus the process is not weakly stationary. Alternatively

$$\begin{aligned}\mathsf{C}[X(t),X(s)] &= \mathsf{C}[Y,Y]\sin(t)\sin(s)+\mathsf{C}[Z,Z] = \sigma_y^2\sin(t)\sin(s)+\sigma_z^2\\ &= \frac{1}{2}\sigma_y^2(\cos(t-s)-\cos(t+s))+\sigma_z^2,\end{aligned}$$

depending on $t-s$ and $t+s$.

2:8. Option 1 has a positive correlation and Option 2 has a negative correlation at $\tau = 1$. Both have zero correlation at $\tau = 2$. The figure indicates positive correlation at $\tau = 1$, and therefore Option 1 seems to be the best alternative.

2:9. (a) This function has the *necessary* properties of a covariance function stated in Theorem 2.2, but one should note that these conditions are not *sufficient*. That the function actually is a covariance function is shown in Chapter 4, where we study the corresponding spectral density.

(b) This is not a covariance function since it is asymmetric.

(c) This is not a covariance function since it is asymmetric.

(d) This is a covariance function; as shown in Theorem 2.3.

(e) This is not a covariance function since there are $r(\tau) > r(0)$.

2:10. $\mathsf{V}[\widehat{M}(x,y)] = \frac{2}{5} + \frac{8}{25}\left(\frac{1}{2}\right)^{\sqrt{2}} \approx 0.52$. Hint: A few different covariances are needed:

$$\begin{aligned}\mathsf{C}[N(x,y),N(x-1,y)] &= \sigma^2\cdot\rho,\\ \mathsf{C}[N(x-1,y),N(x+1,y)] &= \sigma^2\cdot\rho^2,\\ \mathsf{C}[N(x-1,y),N(x,y-1)] &= \sigma^2\cdot\rho^{\sqrt{2}}.\end{aligned}$$

Count how many covariances similar to these there are (in total 25 including the variances $\mathsf{V}[N(x,y)] = \sigma^2$).

2:11. (a) No, $\mathsf{E}[X(t)] = 0$ is constant but $\mathsf{V}[X(t)] = \sin^2 t$ is not. Note the constant phase.

(b) Yes, the phase is randomly distributed, uniformly between 0 and 2π. Further, $W(t) = \sin(t+Z) = \cos(t+Z-\pi/2)$ is a process of similar form as the process $U(t) = \cos(t+Z)$, which we know is weakly stationary (Theorem 2.3 with $A = 1, f_0 = 1/2\pi$). Both $W(t)$ and $U(t)$ are cosine functions started at a random time point, uniformly distributed over a full period of 2π.

2:12. (a) The standard deviation is $\sqrt{\mathsf{V}[\widehat{m}_N]}$ where

$$\begin{aligned}\mathsf{V}[\widehat{m}_N] &= \frac{1}{N^2}\sum_{k=-N+1}^{N-1}(N-|k|)r_X(k)\\ &= \frac{1}{15^2}(15r_X(0)+2\cdot 14r_X(1))\end{aligned}$$

$$= \frac{1}{15^2}(15 \cdot 3 + 2 \cdot 14 \cdot (-1)) = \frac{17}{15^2} = 0.2749^2.$$

(b) The variance is now

$$\begin{aligned} \mathsf{V}[\widehat{m}_N] &= \mathsf{C}\left[\frac{1}{N}\sum_{k=1}^{N} X(\alpha k), \frac{1}{N}\sum_{l=1}^{N} X(\alpha l)\right] \\ &= \frac{1}{N^2}\Big(\sum_{k=-N+1}^{N-1} (N-|k|) r_X(\alpha k)\Big) \\ &= \frac{1}{N^2}(N r_X(0) + 2(N-1) r_X(\alpha)). \end{aligned}$$

If $\alpha > 1$ then $r_X(\alpha) = 0$ and the measurements are uncorrelated. The standard deviation is $\sqrt{r_X(0)/N} = \sqrt{3/15} = 0.4472$ which is larger than the result in a) where the negative correlation between the measurements reduces the standard deviation in the estimate. Thus, $\alpha = 1$ gives the lowest standard deviation.

2:13. (a) $\{Y_N(n)\}$ and $\{Z_M(n)\}$ are smoothed versions of the original signal. If the signal of interest varies very slowly, we can expect it to be approximately constant in the interval, $[n-N+1, \ldots, n]$, and in that case, the averaging will not do any harm to the deterministic signal, but the noise might be reduced.

(b) A large value of N (or M) means averaging over a large number of terms, which will reduce the noise more. However, we will loose time resolution of the deterministic signal.

(c) Let the covariance function of $\{X(n)\}$ be denoted by $r_X(\tau) = r(\tau)$, and solve the following inequalities:

$$\begin{aligned} \mathsf{V}[Y(n)] &\le 0.1, \\ \mathsf{V}\left[\frac{1}{N}\Big(X(n) + \cdots + X(n-N+1)\Big)\right] &\le 0.1, \\ \frac{1}{N^2}\Big(N r_X(0) + 2(N-1) r_X(1)\Big) &\le 0.1, \\ \frac{1}{N^2}\Big(1.2N + N - 1\Big) &\le 0.1, \end{aligned}$$

and

$$\begin{aligned} \mathsf{V}[Z_N(n)] &\le 0.1, \\ \frac{1}{M^2} M r_X(0) &\le 0.1, \\ \frac{1.2}{M} &\le 0.1, \end{aligned}$$

with solutions $N \ge 22$, $M \ge 12$.

(d) While $\{Y(n)\}$ is a sum of random variables with positive correlation, the construction of $\{Z(n)\}$ takes advantage of the fact that $r_X(\pm 2) = 0$, which makes $\{Z(n)\}$ a sum of uncorrelated variables.

2:14. The variance of the estimator is smallest in (b):

$$\mathsf{V}[\widehat{m}_{(a)}] = \frac{1}{9}(3 + 6\mathrm{e}^{-d}) >$$
$$\mathsf{V}[\widehat{m}_{(b)}] = \frac{1}{9}(3 + 4\mathrm{e}^{-d} + 2\mathrm{e}^{-2d}), \quad d > 0.$$

2:15. We denote the observed values by x_t, $t = 1, \ldots, T$. The variance of the estimator, $\widehat{m}_T = \frac{1}{T}\sum_1^T x_t$, can be approximated with:

$$\mathsf{V}[\widehat{m}_T] \approx \frac{1}{T}\sum_{k=-\infty}^{\infty} r(k) = \frac{1}{T}\left(2\sum_{k=0}^{\infty} r(k) - r(0)\right)$$
$$= \frac{1}{T}\left(\left(2\sum_{k=0}^{\infty} 2^{-k/10}\right) - 1\right) = \frac{1 + 2^{-1/10}}{T\left(1 - 2^{-1/10}\right)}.$$

This gives $T \geq 2887$.

2:16. The variance is $\frac{3\sigma^2}{N}$ and it would have been enough with $N/3$ observations to get the same variance if the processes was white instead of colored.

2:17. Assume the correlation function is $\rho(\tau) = \theta^{|\tau|}$. Then, the variance of the mean value estimator, based on N observations is

$$\mathsf{V}[\widehat{m}_N] \approx \frac{1}{N} \cdot \frac{1+\theta}{1-\theta}$$

according to the approximation (2.12). For uncorrelated variables the variance is $\mathsf{V}[\widehat{m}_n] = 1/n$. Setting the variances equal we get $N/n = (1+\theta)/(1-\theta)$. To find θ, let τ_K be the time lag at which the correlation has dropped off to $1/K$, i.e. $1/K = \theta^{\tau_K}$, leading to $\theta = K^{-1/\tau_K}$, and

$$N/n = (K^{1/\tau_K} + 1)/(K^{1/\tau_K} - 1).$$

2:18. $\mathsf{V}[m_1] = 50$, $\mathsf{V}[m_2] = \frac{1738}{9}$. The estimator m_1 is preferred to m_2.

2:20. By the law of large numbers $n^{-1}\sum_1^n Z_n \to \mathsf{E}[X_1] + Y = Y \neq 0$; thus Z_n is not linearly ergodic. Further, $r_Z(t) = 2$ for $t = 0$ and $r_Z(t) = 1$ for $t \neq 0$; thus, the sufficient condition that $\sum_0^\infty r(t)$ is convergent is not satisfied.

Chapter 3

The Poisson process and its relatives

3:1. It is not weakly stationary. The variance of $Y(t) = X(t) - \lambda t$ is the same as that of $X(t)$,

$$\mathsf{V}[Y(t)] = \mathsf{V}[X(t)] = \mathsf{E}[X(t)] = \lambda t,$$

which depends on t.

3:2. The calculations in Section 3.2.3 show the following interesting property of a Poisson process: If one has observed that there has occurred N events in an observation interval $(0,T]$ then the times at which these N events occurred are distributed over the interval as N independent random variables, uniformly distributed over $(0,T]$. This means that the number of events that occurred in an interval $[a,b]$ with $0 < a < b < T$ has a binomial distribution with parameters N and $(b-a)/T$. The arguments can be repeated for more than one interval, leading to a *multinomial distribution* of events in non-overlapping intervals, with probabilities equal to the relative lengths of the intervals.

We calculate the probability that of the four events in $(0,4]$ there were two in $(0,1]$, two in $(1,2]$ and none in $(2,4]$. The probability is

$$\frac{4!}{2!\,2!\,0!}\,0.25^2\,0.25^2\,0.5^0 = \frac{24}{2\cdot 2\cdot 1}\,0.25^4 \approx 0.0234.$$

3:3. The number of accidents, n_T, over $T = 544 \times 10^6$km is assumed to have a Poisson distribution with expectation λT and variance λT. From the observed number, 67, an estimate of λ is

$$\widehat{\lambda}_T = \frac{67}{544 \times 10^6} = 0.123 \times 10^{-6}.$$

According to Example 3.3 an estimate of the standard deviation of $\widehat{\lambda}_T$ is

$$\mathsf{D}[\widehat{\lambda}_T] = \sqrt{\widehat{\lambda}_T/T} = 0.0150 \times 10^{-6},$$

and an approximative 95% confidence interval is (with $q_{\alpha/2} = 1.96$),

$$\widehat{\lambda}_T \pm 1.96 \times \mathsf{D}[\widehat{\lambda}_T] = (0.123 \pm 0.029) \times 10^{-6}.$$

A realistic answer would be between 10 and 15 fires per 100 million train kilometers.

3:4. The number of background counts N_b during a time span T_b is Poisson with mean and variance equal to $\lambda_b T_b$. The number of background + sample counts N_{b+s} during a time span T_s is Poisson with mean and variance $(\lambda_b + \lambda_s)T_s$.

The background and sample intensities can be estimated by

$$\begin{aligned}\widehat{\lambda}_b &= N_b/T_b,\\ \widehat{\lambda}_s &= N_{b+s}/T_s - N_b/T_b,\end{aligned}$$

and the variance of the sample intestity estimate is

$$\mathsf{V}[\widehat{\lambda}_s] = (\lambda_b + \lambda_s)/T_s + \lambda_b/T_b,$$

since the counts are independent. For a fixed $T_b + T_x = T$ this is minimized for

$$\begin{aligned}T_b &= T\,\frac{\sqrt{\lambda_b}}{\sqrt{\lambda_b} + \sqrt{\lambda_b + \lambda_s}},\\ T_s &= T\,\frac{\sqrt{\lambda_b + \lambda_x}}{\sqrt{\lambda_b} + \sqrt{\lambda_b + \lambda_s}}.\end{aligned}$$

An approximate confidence interval for λ_s is given by

$$\widehat{\lambda}_s \pm q_{\alpha/2}\sqrt{(\widehat{\lambda}_b + \widehat{\lambda}_s)/T_s + \widehat{\lambda}_b/T_b},$$

where $q_{\alpha/2}$ is a quantile in the normal distribution.

3:5. Obviously, the event that $T_1 \le s$ is equivalent to $X(s) \ge 1$. Thus the conditional probability is

$$\begin{aligned}\mathsf{P}(T_1 \le s \mid X(t) = 1) &= \frac{\mathsf{P}(T_1 \le s \text{ and } X(t) = 1)}{\mathsf{P}(X(t) = 1)}\\ &= \frac{\mathsf{P}(X(s) = 1)\mathsf{P}(X(t) - X(s) = 0)}{\mathsf{P}(X(t) = 1)}\\ &= \frac{e^{-\lambda s}(\lambda s)^1/1! \times e^{-\lambda(t-s)}(\lambda(t-s))^0/0!}{e^{-\lambda t}(\lambda t)^1/1!} = \frac{s}{t},\end{aligned}$$

in accordance with the result in Section 3.2.3.

3:6. For an inhomogeneous Poisson process with time-dependent intensity function $\lambda(t) = \gamma t$ the number of events in any interval I has a Poisson distribution with mean $\int_{u\in I} \gamma u\,\mathrm{d}u$. Thus, $X(t)$ has a Poisson distribution with mean

$$\mathsf{E}[X(t)] = \int_0^t \lambda(u)\,\mathrm{d}u = \int_0^t \gamma u\,\mathrm{d}u = \gamma t^2/2,$$

while $X(t)-X(s)$ is Poisson with mean

$$\mathsf{E}[X(t)-X(s)] = \int_s^t \gamma u \,\mathrm{d}u = \gamma(t^2-s^2)/2.$$

Thus,

$$\mathsf{P}(T_1 \le s) = \mathsf{P}(X(s) \ge 1) = 1 - \mathsf{P}(X(s)=0) = 1 - e^{-\gamma s^2/2}.$$

Since the number of events in disjoint intervals are independent we obtain, as in the previous exercise, the following expression for the conditional probability

$$\begin{aligned}\mathsf{P}(T_1 \le s \mid X(t)=1) &= \frac{\mathsf{P}(T_1 \le s \text{ and } X(t)=1)}{\mathsf{P}(X(t)=1)} \\ &= \frac{\mathsf{P}(X(s)=1)\mathsf{P}(X(t)-X(s)=0)}{\mathsf{P}(X(t)=1)} \\ &= \frac{e^{-\gamma s^2/2}(\gamma s^2/2)^1/1! \times e^{-\gamma(t^2-s^2)/2}}{e^{-\gamma t^2/2}(\gamma t^2/2)^1/1!} = \frac{s^2}{t^2}.\end{aligned}$$

3:7. (a) Accident risk is affected by many factors. Some of these factors are directly related to time of the day, for example, the amount and type of illumination on the road. Other factors are not directly related to time, but show a systematic variation with time, for example traffic intensity and average speed.

(b) The average number of accidents between six in the morning and six in the evening is

$$\begin{aligned}\mathsf{E}[X(18)-X(6)] &= \int_6^{18} \lambda(t)\,\mathrm{d}t = 0.001\left[2t - \frac{24}{2\pi}\sin(2\pi t/24)\right]_6^{18} \\ &= 0.001(24+24/\pi) = 0.01636.\end{aligned}$$

If we assume that accidents occur according to an inhomogeneous Poisson process, we get the probability of no accidents to $e^{-0.01636} \approx 0.9838$.

3:8. If the number of raisins in a muffin is Poisson distributed with mean m, then the probability of no raisin is e^{-m}. Thus m has to be at least $-\log 0.05 = 3$. The Poisson assumption requires that raisins do not influence each other, and that seems to be a questionable property for raisins in muffins.

3:9. The germ-grain model is an example of a *marked point process* with the random radius of the attached disc is a "mark" on the point. If the disc sizes are statistically independent of the locations and of each other, the model can be defined as an inhomogeneous Poisson process in $\mathbb{R}^3$, generated as follows.

Assume the radius distribution is continuous with probability density function $p_R(r), r > 0$. If there is a point in the center point process at location $(x,y) \in \mathbb{R}^2$, then we draw a random number r from the distribution $p_R(r)$ and put a (marked) point

at $(x,y,r) \in \mathbb{R}^3$. This procedure will generate a Poisson process in $\mathbb{R}^3$ with intensity $\lambda(x,y,r)$ at location (x,y,r) equal to

$$\lambda(x,y,r) = \lambda \times p_R(r).$$

The number of marked points in a region A in $\mathbb{R}^3$ has a Poisson distribution with mean

$$m(A) = \iiint_A \lambda(x,y,r)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}r. \tag{3.1}$$

We now return to the original problem: find the probability that a given fixed point is not covered by any disk; we calculate the probability that the point $(0,0)$, i.e., the origin, is not covered.

A disk with center at (x,y) and radius r covers the origin if and only if $x^2+y^2 \le r^2$. The inequality $r \ge \sqrt{x^2+y^2}$ defines a region A in $\mathbb{R}^3$, and we let $N(A)$ be the number of marked points (x,y,r) that fulfil this criterion. We seek the probability $\mathsf{P}(N(A) \ge 1) = 1 - \mathsf{P}(N(A) = 0)$ that there is at least one marked point in A.

The marked point process is an inhomogeneous Poisson process and, by (3.1), the expected number of points in A is

$$\begin{aligned} m(A) &= \int_{r=0}^{\infty} \iint_{x^2+y^2\le r^2} \lambda(x,y,r)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}r \\ &= \lambda \int_{r=0}^{\infty} p_R(r) \left\{ \iint_{x^2+y^2\le r^2} 1\,\mathrm{d}x\,\mathrm{d}y \right\} \mathrm{d}r = \lambda \int_{r=0}^{\infty} p_R(r)\,\pi r^2\,\mathrm{d}r, \end{aligned}$$

since the double integral within brackets is equal to the area ($= \pi r^2$) of the disk with radius r. With the uniform radius distribution, $p_R(r) = 1/a$ for $0 \le r \le a$, and $p_R(r) = 0$ otherwise, we get

$$m(A) = \int_0^a \frac{\lambda \pi r^2}{a}\,\mathrm{d}r = \lambda \frac{\pi a^2}{3}.$$

Hence, the probability that the origin is covered by at least one disk is $\mathsf{P}(N(A) \ge 1) = 1 - \mathsf{P}(N(A) = 0) = e^{-m(A)} = 1 - \exp(-\lambda \pi a^2/3)$.

3:10. First prove that $m(q) = cq$ for all rational q:

$$\begin{aligned} m(s+0) = m(s) + m(0) &\text{ gives } m(0) = 0, \\ m(s-s) = m(s) + m(-s) &\text{ gives } m(-s) = -m(s), \\ m(s+s+\ldots+s) = m(s) + m(s) + \ldots + m(s) &\text{ gives } m(ns) = nm(s), \\ m(ns) = nm(s) &\text{ gives } m\left(\frac{s}{n}\right) = \frac{1}{n} m(s), \\ m\left(k\frac{s}{n}\right) = km\left(\frac{s}{n}\right) &\text{ gives } m\left(\frac{k}{n}s\right) = \frac{k}{n} m(s). \end{aligned}$$

Thus, $m(qs) = qm(s)$ for all rational q; in particular $m(q) = qm(1)$. If $m(s)$ is continuous then $m(t) = tm(1)$ for all real t.

Chapter 4

Spectral representations

4:1. The spectral density is, by its definition, equal to the integral (or Fourier transform)

$$R_X(f) \triangleq \int_{-\infty}^{+\infty} r_X(\tau) \mathrm{e}^{-i2\pi f\tau} \mathrm{d}\tau.$$

Since this integral sometimes is cumbersome to compute we use the Fourier transformation table which you can find in the table of formulas. Using this table, we get

$$R_X(f) = \frac{\alpha}{\alpha^2 + (2\pi f_0 - 2\pi f)^2} + \frac{\alpha}{\alpha^2 + (2\pi f_0 + 2\pi f)^2}.$$

A nice property of the Fourier transform, $G(f) = \int_{-\infty}^{\infty} g(t)\mathrm{e}^{-i2\pi tf}\mathrm{d}t$, is that it equals the Fourier inverse transform, if the function transform is even symmetric, $g(t) = g(-t)$:

$$\begin{aligned}
G(f) &= \int_{-\infty}^{\infty} g(t)\mathrm{e}^{-i2\pi tf}\mathrm{d}t \\
&= \int_{-\infty}^{\infty} g(t) \left(\underbrace{\cos(-2\pi tf)}_{=\cos(2\pi tf),\text{ even function}} + \underbrace{i\sin(-2\pi tf)}_{=-i\sin(2\pi tf),\text{ odd function}} \right) \mathrm{d}t \\
&= \int_{-\infty}^{\infty} g(t)\cos(2\pi tf)\mathrm{d}t \quad - \quad \underbrace{i\int_{-\infty}^{\infty} g(t)\sin(2\pi tf)\mathrm{d}t}_{\text{Convince yourself that this is zero, if } g(t)=g(-t)} \\
&= \int_{-\infty}^{\infty} g(t)\mathrm{e}^{+i2\pi tf}\mathrm{d}t.
\end{aligned}$$

This means, that for even functions (such as the covariance function for a stationary process), we can use the Fourier transformation table in either direction. It also implies that the Fourier transformation of an even function is real valued.

4:2. (a) $R_X(f) = \sqrt{\frac{\pi}{\alpha}} \exp\left(-\frac{(2\pi f)^2}{4\alpha}\right)$, (b) $R_X(f) = \pi \mathrm{e}^{-2\pi|f|}$.

4:3. $R(f) = \alpha_0^2\delta(f) + \frac{\alpha_1^2}{4}\left(\delta(f-f_0) + \delta(f+f_0)\right)$.

4:4. (a) $r(\tau) = \frac{A}{2}e^{-|\tau|}$, (b) $r(\tau) = \frac{A}{4}e^{-2|\tau|}$, (c) $r(\tau) = \frac{A}{3}\left(\frac{1}{2}e^{-|\tau|} - \frac{1}{4}e^{-2|\tau|}\right)$. Hint: Assume the following partial fraction decomposition

$$R_3(f) = \frac{B}{1+(2\pi f)^2} + \frac{C}{4+(2\pi f)^2}.$$

4:5. a – g, b – f, c – e, d – h .

4:6. (a)

$$\begin{aligned} \mathsf{C}[U(t), U(t+\tau)] &= e^{-100|\tau|}, \\ \mathsf{C}[V(t), V(t+\tau)] &= \begin{cases} \frac{\sin 5\pi\tau - \sin 4\pi\tau}{\pi\tau} & \tau \neq 0, \\ 1 & \tau = 0 \end{cases}. \end{aligned}$$

(b) $U(t) \leftrightarrow$ (a) and $V(t) \leftrightarrow$ (b).

4:7. Only A, E, I can be covariance functions; they have their largest value at $\tau = 0$. G, H, C are spectral densities as they are the only ones that are positive for all values. Consequently D, B, F are realizations. B-A-G belong to the same process as a slowly varying process has a covariance function that varies slowly and a spectral density concentrated at low frequency. D-C-I belong to a faster varying process with higher frequency. F-E-H belong to a 'noisy' process with many frequencies in the spectral density. The covariance function does also decay faster (compare with the spectral density for white noise which is constant and has a covariance function which is zero for all values except for $\tau = 0$).

4:8. Hint: Investigate if there exist any minima for $R(f)$ and investigate the endpoints $f = 0$ and $f = \infty$. As $R(0) = 1 + B/4$, $B \geq -4$. For $f \to \infty$, $R(f) = \frac{1}{f^2}(1/(1/f^2+1) + B/(4/f^2+1))$ can be studied. Then $B \geq -1$. Conclusion $B \geq -1$.

4:9. (a) Covariance functions: A and F, as $r(0) > r(\tau)$ for $\tau > 0$. Spectral densities: C and E, as $R(f) > 0$. Conclusively, realizations: B and D.

(b) B $\leftrightarrow$ A $\leftrightarrow$ E, as the frequencies of the covariance function and the process realization are in concordance with the frequency values of the spectrum, (continuous). D $\leftrightarrow$ C $\leftrightarrow$ F, where the process and the covariance function switches sign for every sample. This is the highest possible frequency of a discrete time process, i.e., $f = 0.5$, where also the spectrum has its maximum.

(c) $h = 3$ for B $\leftrightarrow$ A $\leftrightarrow$ E and $h = 1$ for D $\leftrightarrow$ C $\leftrightarrow$ F.

4:10. (a) See Figure 4.1.

(b) At least twice every second.

(c) For (1),

$$R_{X_s}(f) = \sum_{k=-\infty}^{\infty} R_X(f + kf_s) = \left(\text{for } -\frac{1}{2} < f \leq \frac{1}{2} \right)$$

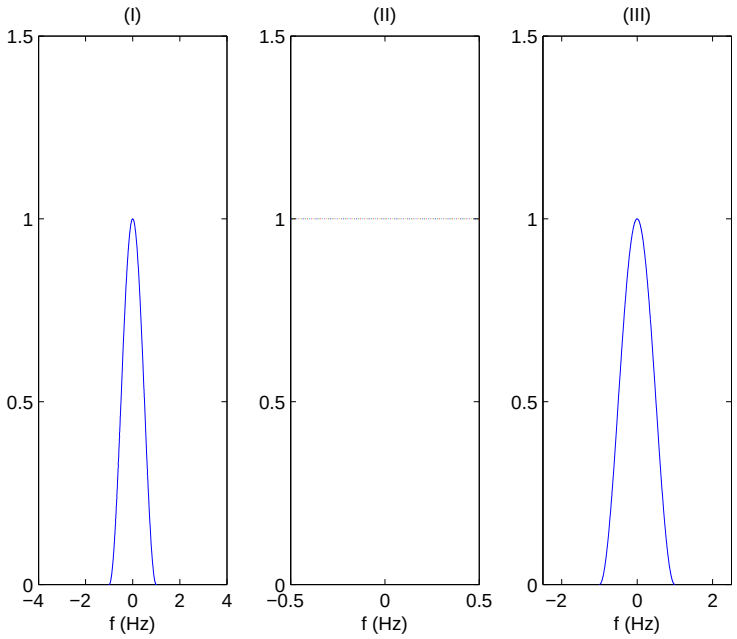

Figure 4.1: Spectral densities of continuous process (I), the process sampled once every second (II) and five times every second (III).

$$
\begin{aligned}
&= \begin{cases} \cos^2(f\cdot\frac{\pi}{2})+\cos^2((f-1)\frac{\pi}{2}),\ f\geq 0, \\ \cos^2(f\cdot\frac{\pi}{2})+\cos^2((f+1)\frac{\pi}{2}),\ f\leq 0, \end{cases} \\
&= \begin{cases} \cos^2(f\cdot\frac{\pi}{2})+\sin^2(f\cdot\frac{\pi}{2})=1, & f\geq 0, \\ \sin^2((f+1)\frac{\pi}{2})+\cos^2((f+1)\frac{\pi}{2})=,1 & f\leq,0 \end{cases} \\
&= 1 \text{ for all } |f|\leq\frac{1}{2};
\end{aligned}
$$

see Figure 4.1.

For (2), the spectrum is $R_X(f),\ |f|\leq 2.5$ Hz; see Figure 4.1.

4:11. (a) The sample frequency should be chosen at least to $f_s = 1/d = 40$ Hz to avoid aliasing.

(b) After sampling the signal contains the frequency area $10\leq|f|\leq 15$ and the disturbance is located at the frequency 10 Hz.

Chapter 5

Gaussian processes

5:1. (a) $1-\Phi(2) \approx 0.02275$.

(b) $1-\Phi\left(\frac{2}{\sqrt{3}}\right) \approx 0.1241$.

(c) $1-\Phi\left(\frac{2}{\sqrt{2}}\right) \approx 0.07865$.

5:2. As usual, it is a good idea to start with the definition of the covariance function:

$$\begin{aligned} r_Y(s,t) &= \mathsf{C}[Y(s),\, Y(t)] = \mathsf{C}[X(s)-0.4X(s-2),\, X(t)-0.4X(t-2)] \\ &= r_X(t-s)+0.16r_X(t-s)-0.4r_X(t-s-2)-0.4r_X(t-s+2). \end{aligned}$$

Since this does only depend on $t-s$ (convince yourself that $r_Y(s+c,t+c) = r_Y(s,t)$ for any constant c) and since $m_Y(t)$ is constant, $\{Y(t)\}$ is a weakly stationary process. It is a Gaussian process, since every process that is a linear combination of Gaussian processes is a Gaussian process. A (real valued) Gaussian weakly stationary process is also strictly stationary.

5:3. $\Phi\left(-\frac{1}{\sqrt{0.6}}\right) \approx 0.0985$. One can imagine the process as a sum of two independent processes $r_X(\tau) = r_A(\tau) + r_B(\tau)$, one process with the covariance function $r_A(\tau) = \frac{1}{1+\tau^2}$ (this is a common process), and one with the covariance function $r_B(\tau) = 1$ (this process is a constant, but a random one!). To simulate a realization of this process, take the outcome of a Gaussian random variable and then add it to a realization of the process with covariance function $r_A(\tau)$.

5:4. $RC \approx 1.92$. Hint: First, you should understand that in order to compute $\mathsf{P}(Z(t)^2 < 1)$ you need the distribution of the random variable $Z(t)^2$. Second, you recognize that this is *not* a Gaussian random variable (show it!) and that the distribution would be tiresome to find. Luckily though, we realize that the event that $Z(t)^2 < 1$ is the same event as $-1 < Z(t) < 1$, and therefore $\mathsf{P}\left(Z(t)^2 < 1\right) = \mathsf{P}\left(-1 < Z(t) < 1\right)$.

5:5. (a) It is found that $\mathsf{E}[Y_{n+1}] = \mathsf{E}[Y_n] + \log a + \mathsf{E}[e_n] = \mathsf{E}[Y_n] + \log a$. As $Y_0 = 0$, we get

$$\mathsf{E}[Y_n] = n \log a.$$

Further

$$\mathsf{C}[Y_m, Y_n] = \mathsf{C}[Y_m, Y_{n-1}] + \mathsf{C}[Y_m, e_{n-1}] = \mathsf{C}[Y_m, Y_{n-1}].$$

for $m < n$. Repeating,

$$\mathsf{C}[Y_m, Y_n] = \mathsf{C}[Y_m, Y_m] = \mathsf{V}[Y_m]$$

if $m < n$. On the other hand

$$\mathsf{V}[Y_n] = \mathsf{V}[Y_{n-1}] + \mathsf{V}[e_{n-1}] = \mathsf{V}[Y_{n-1}] + (K \log a)^2 = n(K \log a)^2,$$

giving

$$\mathsf{C}[Y_m, Y_n] = \min(m,n)(K \log a)^2,$$

which gives a non-stationary process.

(b) The probability is

$$P((X_{n+1} - X_n)/X_n > 0.038) = P((X_{n+1} - X_n)/X_n > 100p) =$$

$$P(X_{n+1}/X_n > 1 + 100p) = P(\log X_{n+1} - \log X_n > \log(1 + 100p)) =$$

$$P(Y_{n+1} - Y_n > \log(1 + 100p)) = P(e_n > \log(1 + 100p) - \log a) =$$

$$1 - \Phi\left(\frac{\log(1 + 100p) - \log a - 0}{K \log a}\right) \approx 1 - \Phi(99/30) = 0.00048.$$

In average $250 \cdot 0.00048 = 0.12$ occasions per year.

5:6. (a) The Wiener process is a Gaussian process, and hence, $Y(t)$ is Gaussian distributed. $Z(t)$ is a linear combination of Gaussian random variables, and is thus Gaussian. It remains to compute the expectation and variance. The expectation of $Z(t)$ is

$$\begin{aligned} \mathsf{E}[Z(t)] &= \mathsf{E}\left[\frac{Y(t) - Y(t/2)}{\sqrt{t}}\right] \\ &= \frac{\mathsf{E}[Y(t)] - \mathsf{E}[Y(t/2)]}{\sqrt{t}} = 0, \end{aligned}$$

and the variance is

$$\begin{aligned} \mathsf{V}[Z(t)] &= \mathsf{C}\left[\frac{Y(t) - Y(t/2)}{\sqrt{t}}, \frac{Y(t) - Y(t/2)}{\sqrt{t}}\right] \\ &= \frac{1}{t}(t - t/2 - t/2 + t/2) = \frac{1}{2}. \end{aligned}$$

We conclude that $Z(t) \sim N\left(0, \frac{1}{2}\right)$.

(b) The expectation (found in (a)) does not depend on the time, t. We need to compute the covariance function:

$$r_Z(s,t) \triangleq \mathsf{C}[Z(s),\, Z(t)] = \mathsf{C}\left[\frac{Y(s) - Y(s/2)}{\sqrt{s}}, \frac{Y(t) - Y(t/2)}{\sqrt{t}}\right]$$

$$= \frac{1}{\sqrt{ts}}\left(\min(s,t) - \min(s,t/2) - \min(s/2,t) + \min(s/2,t/2)\right)$$

$$= \begin{cases} \frac{1}{\sqrt{ts}}(s - s - s/2 + s/2) & = 0 & s \le t/2 \\ \frac{1}{\sqrt{ts}}(s - t/2 - s/2 + s/2) & = \frac{1}{\sqrt{ts}}(s - t/2) & t/2 \le s \le t \\ \frac{1}{\sqrt{ts}}(t - t/2 - s/2 + t/2) & = \frac{1}{\sqrt{ts}}(t - s/2) & t \le s \le 2t \\ 0 & & s > 2t \end{cases}.$$

Since this is not a function of the distance between s and t, we conclude that the process is not weakly stationary.

5:7. $r_B(s,t) = \min(s,t) - st$.

5:8. $m = 10^4 \cdot \int_0^\infty 10^{-2} e^{-t}\,\mathrm{d}t = 100$. For $\tau \ge 0$,

$$\begin{aligned} r(\tau) &= 10^4 \cdot \int_\tau^\infty 10^{-2} e^{-t} \cdot 10^{-2} \cdot e^{-(t-\tau)}\,\mathrm{d}t = e^{-\tau} \cdot \int_\tau^\infty e^{-2(t-\tau)}\,\mathrm{d}t \\ &= e^{-\tau} \cdot \int_0^\infty e^{-2t}\,\mathrm{d}t = e^{-\tau}/2. \end{aligned}$$

At every time point t, $X(t)$ is determined by many very small contributions (each $\le 10^{-2}$) and they sum to approximately 100. Then $X(t)$ is approximately Gaussian distributed with

$$\begin{aligned} \mathsf{E}[X(t+1) - X(t)] &= 0 \\ \mathsf{V}[X(t+1) - X(t)] &= \mathsf{C}[X(t+1) - X(t), X(t+1) - X(t)] \\ &= \mathsf{V}[X(t+1)] + \mathsf{V}[X(t)] - 2\mathsf{C}[X(t), X(t+1)] \\ &= \frac{1}{2} + \frac{1}{2} - 2 \cdot \frac{1}{2} e^{-1} = 1 - e^{-1} \end{aligned}$$

Conclusion: $X(t+1) - X(t) \sim N\left(0, 1 - e^{-1}\right)$ and therefore

$$\mathsf{P}(X(t+1) - X(t) > 2) = 1 - \Phi\left(\frac{2}{\sqrt{1 - e^{-1}}}\right) = 0.006.$$

Chapter 6

Linear filters – general theory

6:1. With

$$\begin{aligned} Y(t) &= \int h(u)X(t-u)\,\mathrm{d}u = \int (\delta_0(u)-\delta_1(u))X(t-u)\,\mathrm{d}u \\ &= X(t)-X(t-1) \end{aligned}$$

we get

$$\begin{aligned} \mathsf{P}(Y(1) > 3+Y(0)) &= \mathsf{P}(X(1)-X(0) > 3+X(0)-X(-1)) \\ &= \mathsf{P}\underbrace{(X(1)+X(-1)-2X(0)}_{Z} > 3) \\ &= \mathsf{P}(Z>3) = 1-\mathsf{P}(Z\le 3) \\ &= 1-\Phi(\frac{3}{\sqrt{12/5}}) \approx 0.026 \end{aligned}$$

as $\{X(t)\}$ is a Gaussian process and

$$\begin{aligned} \mathsf{E}[Z] &= \mathsf{E}[X(1)]+\mathsf{E}[X(-1)]-2\mathsf{E}[X(0)] = 3+3-2\cdot 3 = 0, \\ \mathsf{V}[Z] &= \mathsf{V}[X(1)]+\mathsf{V}[X(-1)]+(-2)^2\mathsf{V}[X(0)]+2\cdot\mathsf{C}[X(1),X(-1)] \\ &\quad +2\cdot(-2)C[X(1),\,X(0)]+2\cdot 1(-2)C[X(-1),X(0)] \\ &= r_X(0)+r_X(0)+4r_X(0)+2r_X(2)-4r_X(1)-4r_X(-1) \\ &= 1+1+4+2\cdot\frac{1}{5}-4\cdot\frac{1}{2}-4\cdot\frac{1}{2} = \frac{12}{5}. \end{aligned}$$

6:2. The output is a Gaussian process as the filter is linear with Gaussian input. Further, the output spectral density is, for $|f| \le q$,

$$R_Y(f) = |H(f)|^2 R_X(f) = \left|\frac{2}{1+|f|}\right|^2 (1+|f|) = \frac{4}{1+|f|}.$$

The variance is

$$r_Y(0) = 2\int_0^1 \frac{4}{1+f}\,\mathrm{d}f = 8\,[\ln(1+f)]_0^1 = 8\ln 2 \approx 5.545.$$

The expected value is

$$m_Y = H(0)m_X = 2\cdot 1 = 2,$$

giving

$$\mathsf{P}\,(Y(t) \le 4) = \Phi\left(\frac{4-2}{\sqrt{8\ln(2)}}\right) \approx \Phi(0.85) \approx 0.802.$$

6:3. The covariance function of the output signal is $r_Y(\tau) = -\frac{2}{3}\mathrm{e}^{-|\tau|} + \frac{4}{3}\mathrm{e}^{-2|\tau|}$.

6:4. (a) The filter is causal with impulse response

$$h(u) = \begin{cases} 1 & \text{for } 0 \le u \le 2, \\ 0 & \text{otherwise.} \end{cases}$$

(b) The frequency function of the filter is

$$H(f) = \begin{cases} \frac{1-\mathrm{e}^{-i4\pi f}}{i2\pi f} & \text{for } f \neq 0, \\ 2 & \text{for } f = 0. \end{cases}$$

(c) The output spectral density is

$$R_Y(f) = \begin{cases} 4 & \text{for } f = 0, \\ \frac{4(1-\cos 2\pi f)(1-\cos 4\pi f)}{(2\pi f)^4} & \text{for } f \neq 0. \end{cases}$$

6:5. (a) $\mathsf{E}[Y(t)^2] = 2\int_{f_0-\Delta f/2}^{f_0+\Delta f/2} R_X(f)\,\mathrm{d}f$.

(b) $R_X(f_0) \approx \frac{1}{2\Delta f}\mathsf{E}[Y(t)^2]$.

6:6. (a) The filter frequency function is

$$H(f) = 1 \quad \text{for } |f| \le 10,$$

and zero for all other values. This means that the frequencies $f_1 = 4.5$ Hz and $f_2 = 7$ Hz go unchanged through the filter while $f_3 = 11$ Hz is completely stopped.

(b) By sampling with $d = 0.1$ the sampling frequency is $f_s = 1/d = 10$ Hz and therefore the highest possible frequency is $f_s/2 = 5$ Hz. The frequency $f_1^s = 4.5 \cdot 0.1 = 0.45$ after sampling but $f_2 = 7$ Hz is aliased to $10 - 7 = 3$ Hz giving $f_2^s = 3 \cdot 0.1 = 0.3$.

6:7. The spectral density of the derivative is $(2\pi f)^2$ times the spectral density of the process:

$$R_{X'}(f) = (2\pi f)^2 R_X(f) = \pi(2\pi f)^2 \mathrm{e}^{-2\pi|f|}.$$

The covariance function of the derivative is the negated second derivative of the covariance function. The covariance function is the inverse Fourier transformation of the spectral density. The table of Fourier transforms gives us that $\underset{\tau\to f}{\mathscr{F}}\left(\mathrm{e}^{-\alpha|\tau|}\right) = \frac{2\alpha}{\alpha^2+(2\pi f)^2}$.

Since the function is even symmetric, the Fourier transform of $\mathrm{e}^{-\alpha|\tau|}$ is equal to the inverse Fourier transformation of $\mathrm{e}^{-\alpha|\tau|}$. Thus,

$$\begin{aligned} r_X(\tau) &= \underset{f\to\tau}{\mathscr{F}}^{-1} R_X(f) = \underset{f\to\tau}{\mathscr{F}}^{-1} \pi \mathrm{e}^{-2\pi|f|} \\ &= \underset{f\to\tau}{\mathscr{F}} \pi \mathrm{e}^{-2\pi|f|} = \pi \underset{f\to\tau}{\mathscr{F}} \mathrm{e}^{-2\pi|f|} \\ &= \pi \frac{2\cdot 2\pi}{(2\pi)^2 + (2\pi\tau)^2} = \frac{1}{1+\tau^2} \end{aligned}$$

This gives the covariance function of the derivative,

$$r_{X'}(\tau) = -\frac{d^2 r_X(\tau)}{d\tau^2} = -\frac{d^2}{d\tau^2}\frac{1}{1+\tau^2} = \frac{2-6\tau^2}{(1+\tau^2)^3}.$$

6:8. A function $r(t)$ is a covariance function of a (real valued) stationary process if and only if its Fourier transform, $R(f)$, is (even symmetric), non negative, and integrable with

$$\int_{-\infty}^{\infty} R(f)\,\mathrm{d}f < \infty.$$

We compute the Fourier transform of $r_X(t)$ using the table of Fourier transforms:

$$\begin{aligned} R_X(f) &= \underset{t\to\tau}{\mathscr{F}} r_X(t) = 4\frac{2\cdot\frac{1}{4}}{\frac{1}{4^2}+(2\pi f)^2} - 2\frac{2\cdot\frac{1}{2}}{\frac{1}{2^2}+(2\pi f)^2} \\ &= \frac{24}{(1+16(2\pi f)^2)(1+4(2\pi f)^2)}. \end{aligned}$$

This function is non negative for all $f \in \mathbb{R}$ and it behaves as $1/f^4$ when $f \to \pm\infty$ and is thus integrable with $\int R(f)\,\mathrm{d}f < \infty$. A stationary process is differentiable if and only if $\int (2\pi f)^2 R(f)\,\mathrm{d}f < \infty$. It is equivalent to that the covariance function is twice differentiable. However; in this case it is more difficult to show that the covariance function is twice differentiable, since the two terms $4\mathrm{e}^{-2\pi|t|/4}$ and $-2\mathrm{e}^{-|t|/2}$ are not differentiable at $t = 0$.

As mentioned, another possibility is to consider the integral

$$\int_{-\infty}^{\infty} (2\pi f)^2 R(f)\,\mathrm{d}f = \int_{-\infty}^{\infty} \frac{24(2\pi f)^2}{(1+16(2\pi f)^2)(1+4(2\pi f)^2)}\,\mathrm{d}f.$$

The integrand is of the order $1/f^2$ when $f \to \pm\infty$, and we conclude that $\int (2\pi f)^2 R(f)\,\mathrm{d}f < \infty$.

6:9.

$$r_{Y,Z}(n,n+k) \;=\; \mathsf{C}[Y(n),Z(n+k)] = \begin{cases} \sigma^2, & k=0, \\ -\sigma^2, & k=1, \\ \sigma^2, & k=2, \\ 3\sigma^2, & k=-1, \\ 0, & \text{otherwise.} \end{cases}$$

6:10. (a) The cross-covariance function is $r_{X,Y}(\tau) = \int_0^\infty h(s) r_X(\tau - s)\,\mathrm{d}s = (h * r_X)(\tau)$, where $r_X(\tau)$ is the covariance function of $\{X(t)\}$ and $*$ denotes convolution.

The cross-spectral density is $R_{X,Y}(f) = H(f) R_X(f)$, where $H(f)$ is the frequency response of the filter and $R_X(f)$ is the spectral density of $\{X(t)\}$.

(b) The cross-covariance function is $r_{X,Y}(\tau) = \sigma^2 h(\tau)$, where σ^2 is the variance of the white noise and the cross-spectral density is $R_{X,Y}(f) = \sigma^2 H(f)$.

Chapter 7

AR, MA, and ARMA-models

7:1. Answer: $X(t) - \frac{1}{2}X(t-1) = e(t)$, $\mathsf{V}[e(t)] = 7.5$.

The Yule-Walker equations give

$$\begin{cases} r(0) + ar(1) = \sigma^2 \\ r(1) + ar(0) = 0 \end{cases} \iff \begin{cases} 10 + a \cdot 5 = \sigma^2 \\ 5 + a \cdot 10 = 0 \end{cases}$$

as $r(0) = 10, r(1) = 5$

7:2. The Yule-Walker equations give

$$\begin{aligned} r(0) - r(1) + 0.5r(2) &= 4, \\ r(1) - r(0) + 0.5r(1) &= 0, \\ r(2) - r(1) + 0.5r(0) &= 0, \end{aligned}$$

with the solution $r(0) = 9.6$, $r(1) = 6.4$, and $r(2) = 1.6$. The spectral density is

$$\begin{aligned} R(f) &= \frac{\sigma^2}{|\sum_k a_k \mathrm{e}^{-i2\pi fk}|^2} = \frac{4}{|1 - \mathrm{e}^{-i2\pi f} + 0.5\mathrm{e}^{-i4\pi f}|^2} \\ &= \frac{4}{2.25 - 3\cos 2\pi f + \cos 4\pi f}. \end{aligned}$$

7:3. $\sigma^2 = 1$, $a_2 = 0.5$, and $r_Y(2) = 0.4$.

7:4. $q = 2$, $b_1 = 0$, $b_2 = \frac{1}{2}$, $R_Y(f) = \frac{5}{4} + \cos 4\pi f$, $f \in \left(-\frac{1}{2}, \frac{1}{2}\right]$.

7:5. (a) $\{Z_t\}$ is an MA(3)-process if we choose M such that $\{Z_t\}$ is a zero mean process. This gives $M = \mathsf{E}[X_t] = \mathsf{E}[e_t + 2e_{t-1} - e_{t-3}] = 2m$. Since $\{Z_t\}$ is equivalent to the MA-process

$$Z_t = c_t + 2c_{t-1} - c_{t-3},$$

where c_t is an innovation process with zero mean, we can compute the covariance function of $\{Z_t\}$ by using the standard formula for covariance of an MA-process:

$$r(\tau) = \begin{cases} \sigma^2 \sum_{j-k=\tau} b_j b_k & |\tau| \leq q \\ 0 & |\tau| > q, \end{cases}$$

where $q = 3$, $b_0 = 1$, $b_1 = 2$, $b_2 = 0$, and $b_3 = -1$. This gives

$$r_Z(\tau) = \begin{cases} \sigma^2(1+4+0+1) & \tau = 0, \\ \sigma^2(2+0+0) & |\tau| = 1, \\ \sigma^2(0-2) & |\tau| = 2, \\ \sigma^2(-1) & |\tau| = 3, \\ 0 & |\tau| > 3. \end{cases}$$

The spectral density is computed in the same manner:

$$\begin{aligned} R_Z(f) &= \sigma^2 \left| \sum_{k=0}^{q} b_k \mathrm{e}^{-i2\pi fk} \right|^2 \\ &= \sigma^2 \left((1 + 2\cos 2\pi f - \cos 6\pi f)^2 + (2\sin 2\pi f - \sin 6\pi f)^2 \right) \\ &= \sigma^2 \left(6 + 4\cos 2\pi f - 4\cos 4\pi f - 2\cos 6\pi f \right). \end{aligned}$$

(b) Since $\{X_t\}$ and $\{Z_t\}$ only differ by an additive constant, they will share the same covariance function and spectral density (if you don't find it too obvious, prove it!). The expectation function is $\mathsf{E}[X_t] = 2m$.

7:6. Every covariance function $r(\tau)$ has its maximum in zero: $r(0) \geq r(\tau)$. Thus, Figures (a) and (c) are realizations and (b) and (d) are estimated covariance functions. In each time step, it is likely that a realization of $\{X(t)\}$ changes sign, since it is a realization of an AR(1)-process with positive parameter. This fits well with realization (a) and covariance function (d). The MA-process is a smoothing average of a white noise process, so it will change more slowly, which matches realization (c). Theoretically, the covariance function of $\{Y_t\}$ will be zero for time lags greater or equal to three. This corresponds to the estimated covariance function (b).

7:7. (a) A spectral density function is always positive, which means that only Figures A and F are spectral densities. The covariance function always has the largest value at zero, which holds for D and E. Then B and C have to be realizations.

(b) Figure α illustrates an AR-process of order four as there are four poles. Figure β gives an MA-process of order three as there are three zeros. For the AR-process the spectral density F has two peaks at frequencies 0.1 and 0.3, which correspond to the angle frequencies of the poles $\omega = 2\pi f$.

The spectral density A might come from the MA-process as it has low energy corresponding to the zeros. The covariance function E has values that differ from zero for large τ, so it has to be the AR-process. Covariance function D seems to be zero at $\tau = 4$

and above, so it should be an MA(3)-process. The realization C has higher frequency than B. Accordingly, realization C must belong to the spectral density A and the MA(3)-process, while realization B belongs to the spectral density F and the AR-process. In summary: α – AR(4) – F-E-B, and β – MA(3) – A-D-C.

7:8. C belongs to I and 1 as a low frequency for the spectral density means a smaller angle for the pole and a slower variation for the covariance function. B belongs to 2 and III as a pole far from the unit circle gives a damped covariance function and a peak in the spectral density which is not that sharp. A belongs to II and 3 where a higher frequency in the spectral density means a larger angle for the pole and a faster variation of the covariance function.

7:9. $(1-1.559+0.81)\cdot \mathsf{E}[W_t] = 3.5 \Rightarrow \mathsf{E}[W_t] = \frac{3.5}{0.251} = 13.94$. $W_t - 13.94$ is an AR(2)-process and $\mathsf{V}[W_t] = r(0)$ is given from the Yule-Walker equations:

$$\begin{cases} r(0) - 1.559\, r(1) + 0.81\, r(2) = 4 \\ r(1) - 1.559\, r(0) + 0.81\, r(1) = 0 \\ r(2) - 1.559\, r(1) + 0.81\, r(0) = 0 \end{cases}$$

Solution: $r(0) = 45.062, r(1) = 38.813, r(2) = 24.009$, where the standard deviation is given as $\sqrt{\mathsf{V}[W_t]} = \sqrt{r(0)} \approx 6.71$ m.

7:10. The model for the sampled process is

$$Y_m = X_{2m} = -a_1 X_{2m-1} + e_{2m},$$

where

$$X_{2m-1} = -a_1 X_{2m-2} + e_{2m-1} = -a_1 Y_{m-1} + e_{2m-1}$$

and

$$Y_m = a_1^2 Y_{m-1} + \underbrace{-a_1 e_{2m-1} + e_{2m}}_{u_m}$$

Here u_m has $\mathsf{V}[u_m] = (a_1^2+1)\sigma^2$, and is uncorrelated with $Y_{m-1}, Y_{m-2}, \ldots$ Conclusion: $Y_m - a_1^2 Y_{m-1} = u_m$ is an AR(1)-process with parameter $-a_1^2$ and $\mathsf{V}(u_m) = (a_1^2+1)\sigma^2$.

7:11.

$$\mathsf{C}[X_{t+k},\, e_t] = \begin{cases} 0, & k < 0, \\ 2\cdot 0.4^k, & k \geq 0. \end{cases}$$

7:12. (a) $\widehat{r}_X(0) \approx 0.64$, $\widehat{r}_X(1) \approx -0.12$.

(b) $\widehat{a} \approx 0.19$, $\widehat{(\sigma^2)} \approx 0.62$.

7:13. We use the following estimator:

$$\widehat{m} = \frac{1}{10}\sum_{t=1}^{10} X_t \quad .$$

To compute the variance we need the covariance function of the process $\{X_t\}$, which is the same as the covariance function of the process $\{X_t - m\}$. We compute the covariance function using the Yule-Walker equations:

$$\begin{aligned} r_X(k) + 0.25 r_X(k-1) &= 0 \quad \text{for } k > 0, \\ r_X(0) + 0.25 r_X(1) &= \sigma^2 = 1. \end{aligned}$$

This gives

$$r_X(0) = \frac{16}{15}, \qquad r_X(k) = (-0.25)^k \frac{16}{15}, \quad k > 0.$$

The variance of the estimator is

$$\begin{aligned} \mathsf{V}[\widehat{m}] &= \mathsf{V}\left[\frac{1}{10}\sum_{t=1}^{10} X_t\right] = \frac{1}{100}\mathsf{C}\left[\sum_{t=1}^{10} X_t, \sum_{s=1}^{10} X_s\right] \\ &= \frac{1}{100}\Big(10 r_X(0) + 2\cdot 9 r_X(1) + 2\cdot 8 r_X(2) + \cdots + 2 r_X(9)\Big) \\ &= \frac{1}{100}\frac{1}{15}\left(10\cdot 16 - 18\cdot 4 + \cdots - 2\cdot\frac{1}{16384}\right) \approx 0.06741. \end{aligned}$$

The expectation of $\widehat{m}$ is m, and $\widehat{m}$ is Gaussian since it is a linear combination of the Gaussian random variables e_t. Thus $\widehat{m} \sim N(m, 0.06741)$. A point estimate of m is $10^{-1}\sum_{t=1}^{10} x_t = 0.7347$. A 95% confidence interval is given by: $0.7347 \pm 1.96 \cdot \sqrt{0.06741} = (0.2258,\ 1.2436)$.

7:14. With $A(T^{-1}) = 1 - T^{-1}$, $C(T^{-1}) = 1 + T^{-1}$, and

$$e_t = \frac{A(T^{-1})}{C(T^{-1})} X_t,$$

we get

$$\begin{aligned} X_{t+1} &= \frac{C(T^{-1})}{A(T^{-1})} e_{t+1} = e_{t+1} + \frac{C(T^{-1}) - A(T^{-1})}{A(T^{-1})T^{-1}} T^{-1} e_{t+1} \\ &= e_{t+1} + \frac{1}{C(T^{-1})} X_t = \sum_{k=0}^{\infty} (-0.5)^k X_{t-k}. \end{aligned}$$

7:15. From the Yule-Walker equation with $k > 0$,

$$r_X(k) = -(a_1 r_X(k-1) + a_2 r_X(k-2) + a_3 r_X(k-3))$$

we get $r_X(4) = 0.0244$, $r_X(5) = 0.0012$.

Chapter 8

Linear filters – applications

8:1. Draw a picture to illustrate the symmetry/non-symmetry.

8:2. The Wiener filter is given by

$$H(f) = \frac{R_s(f)}{R_S(f) + R_N(f)},$$

where

$$R_S(f) = \sum_\tau r_S(\tau) e^{-i2\pi f\tau} = 2 - 2\cos(2\pi f),$$

and

$$R_N(f) = \sum_\tau r_N(\tau) e^{-i2\pi f\tau} = 2 + 2\cos(2\pi f).$$

We get

$$H(f) = 0.5 - 0.5\cos(2\pi f) = 0.5 - 0.25(e^{-i2\pi f} + e^{i2\pi f})$$

for $-0.5 < f \leq 0.5$. The impulse response is $h(0) = 0.5$, $h(\pm 1) = -0.25$ and zero for all other u.

8:3. The spectral density $R_X(f)$ of the temperature and the spectral density $R_N(f)$ of the noise are given by (use the Fourier transformation table!)

$$R_X(f) = \frac{4}{4 + (2\pi f)^2}, \qquad R_N(f) = \frac{40}{20^2 + (2\pi f)^2}.$$

The optimal filter is given by

$$H(f) = \frac{R_X(f)}{R_X(f) + R_N(f)},$$

where $H(f)$ is the frequency response of the filter.

The covariance function of the output is most easily found via the spectral density. It is

$$R_Y(f) = |H(f)|^2 (R_X(f) + R_N(f))$$

$$
\begin{aligned}
&= \frac{R_X^2(f)}{R_X(f)+R_N(f)} = \frac{\left(\frac{4}{4+(2\pi f)^2}\right)^2}{\frac{4}{4+(2\pi f)^2}+\frac{40}{400+(2\pi f)^2}} \\
&= \frac{16(400+(2\pi f)^2)}{(4+(2\pi f)^2)(1760+44(2\pi f)^2)} = \frac{4}{11}\frac{(400+(2\pi f)^2)}{(4+(2\pi f)^2)(40+(2\pi f)^2)} \\
&= \frac{4}{11}\left(\frac{11}{(4+(2\pi f)^2)}-\frac{10}{(40+(2\pi f)^2)}\right),
\end{aligned}
$$

where partial fraction decomposition is applied. The covariance function is

$$r_Y(\tau) = \mathrm{e}^{-2|\tau|} - \frac{\sqrt{10}}{11}\mathrm{e}^{-2\sqrt{10}|\tau|}.$$

8:4. (a)

$$
\begin{aligned}
H(f) &= \begin{cases} \frac{1}{1+\frac{100}{|f|}}, & 100 \le |f| \le 1000, \\ 0, & \text{otherwise}, \end{cases} \\
\mathrm{SNR}_{\max} &\approx 5.28.
\end{aligned}
$$

(b) $\mathrm{SNR}_{\text{engineer}} \approx 3.91$.

8:5. From the Yule-Walker equations, $a_1 = -3/4$, $a_2 = 1/2$, and $\sigma^2 = 9/4$.

Since the white noise has constant spectral density the Wiener filter frequency function is

$$
\begin{aligned}
H(f) &= \frac{9/4}{9/4+29/16-9/4\cos(2\pi f)+\cos(4\pi f)} \\
&= \frac{1}{65/36-\cos(2\pi f)+4/9\cos(4\pi f)}.
\end{aligned}
$$

8:6. (a) The filter impulse response is, for a constant c,

$$
\begin{aligned}
h(u) &= \begin{cases} cA\mathrm{e}^{-b(T-u)}, & t \le T, \\ 0, & \text{otherwise}, \end{cases} \\
\mathrm{SNR}_{\max} &= \frac{A^2}{2bN_0}.
\end{aligned}
$$

(b)

$$
\begin{aligned}
h_1(t) &= \begin{cases} cA\mathrm{e}^{-b(T-t)}, & 0 \le t \le T, \\ 0, & \text{otherwise}. \end{cases} \\
\mathrm{SNR}_{h_1} &= \frac{A^2\left(1-\mathrm{e}^{-2bT}\right)}{2bN_0}.
\end{aligned}
$$

(c) $T > \frac{\ln 100}{2b}$.

8:7.

$$h(t) = \begin{cases} \frac{T-t}{\varepsilon} & T-\varepsilon \le t \le T \\ 1 & \varepsilon \le t \le T-\varepsilon \\ \frac{t}{\varepsilon} & 0 \le t \le \varepsilon \\ 0 & \text{otherwise.} \end{cases}$$

According to (8.25), the optimal signal-to-noise ratio is

$$\text{SNR} = \frac{1}{R_0}\left(T - \frac{4}{3}\varepsilon\right).$$

8:8. (a)

$$\begin{aligned} &\mathsf{P}\,(\text{Suggest that the signal is present} \mid \text{no signal is present}) \\ &= 1 - \Phi\left(\frac{1}{\sqrt{2}}\right) \approx 0.240. \end{aligned}$$

(b)

$$h(k) = \begin{cases} 1, \ k = T-3, \ T-2, \ T-1, \ T, \\ 0, \ \text{otherwise.} \end{cases}$$

This filter is causal if, for example, the decision time, T, is 3. The decision level is 2.

With symmetric error probabilities, we obtain

$$\begin{aligned} &\mathsf{P}\,(\text{Suggest that the signal is present} \mid \text{no signal present}) \\ &= \mathsf{P}\,(\text{Suggest that no signal is present} \mid \text{signal is present}) \\ &= 1 - \Phi(1) \approx 0.1587. \end{aligned}$$

8:9. For the given impulse response,

$$\begin{aligned} &\min_{a,b} \mathsf{E}[(aX(t) + bX(t+1) - S(t))^2] \\ &\quad = \min_{a,b} \mathsf{E}[(aS(t) + 0.5aS(t-1) + aN(t) + bS(t+1) \\ &\qquad\qquad + 0.5bS(t) + bN(t+1) - S(t))^2] \\ &\quad = \min_{a,b}((a-1+0.5b)^2 + 0.25a^2 + b^2)r_S(0) \\ &\qquad\qquad + (a^2+b^2)r_N(0) + 2(0.5a(a-1+0.5b) \\ &\qquad\qquad + b(a-1+0.5b))r_S(1) + 2abr_N(1)). \end{aligned}$$

We get

$$\min f(a,b) = \frac{7}{2}a^2 + \frac{7}{2}b^2 + \frac{3}{2}ab - 3a + 2$$

and differentiation with respect to a, b gives

$$\begin{aligned}\frac{\partial f}{\partial a} &= 7a + \frac{3b}{2} - 3 = 0,\\ \frac{\partial f}{\partial b} &= 7b + \frac{3}{2}a = 0,\end{aligned}$$

giving $b = -18/187 \approx -0.0963$ and $a = 84/187 \approx 0.449$ and the impulse response $h(0) = 0.449$, $h(-1) = -0.0963$ and zero for all other t.

8:10. First initialize the state variable using the unconditional stationary distribution of an AR(1)-process,

$$X_0 \sim N(0, \sigma_e^2/(1-\phi_1^2)),$$

with variance

$$V_{XX}(0 \mid 0) = \frac{\sigma_e^2}{1-\phi_1^2},$$

and take

$$\widehat{X}_{0|0} = X_0.$$

Then, for $t = 0, 1, 2, \ldots$, repeate the following steps.

Predict:

$$\begin{aligned}\widehat{X}_{t+1|t} &= \phi_1 \widehat{X}_{t|t},\\ V_{XXX}(t+1 \mid t) &= \phi_1^2 + V_{XX}(t \mid t) + \sigma_e^2,\\ \widehat{Y}_{t+1|t} &= \widehat{X}_{t+1|t}.\end{aligned}$$

Update:

$$\begin{aligned}V_{YY}(t+1|t) &= V_{XX}(t+1 \mid t) + \sigma_f^2,\\ C_{XY}(t+1 \mid t) &= X_{XX}(t+1 \mid t),\\ K(t+1) &= \frac{V_{XX}(t+1 \mid t)}{V_{XX}(t+1 \mid t) + \sigma_f^2},\\ \widehat{X}_{t+1|t+1} &= \widehat{X}_{t+1|t} + K(t+1)(Y_{t+1} - \widehat{Y}_{t+1|t}),\\ V_{XX}(t+1 \mid t+1) &= \frac{\phi_1^2 V_{XX}(t \mid t) + \sigma_e^2)\sigma_f^2}{\phi_1^2 V_{XX}(t \mid t) + \sigma_e^2 + \sigma_f^2}.\end{aligned}$$

Chapter 9

Frequency analysis and spectral estimation

9:1.

$$\begin{aligned}\mathscr{X}(f+1) &= \sum_{t=0}^{n-1} x(t)e^{-i2\pi(f+1)t}, \\ &= \sum_{t=0}^{n-1} x(t)e^{-i2\pi ft}e^{-i2\pi t} = \sum_{t=0}^{n-1} x(t)e^{-i2\pi ft},\end{aligned}$$

since $e^{-i2\pi t} = 1$ for all integer values of t.

9:2.

$$\begin{aligned}\mathsf{E}[(\widehat{R}_x(f) - R_x(f))^2] &= \mathsf{E}[(\widehat{R}_x(f))^2] - 2\mathsf{E}[\widehat{R}_x(f)]R_x(f) + R_x^2(f) \\ &= \mathsf{E}[(\widehat{R}_x(f))^2] - \mathsf{E}^2[\widehat{R}_x(f)] \\ &\qquad + \mathsf{E}^2[\widehat{R}_x(f)] - 2\mathsf{E}[\widehat{R}_x(f)]R_x(f) + R_x^2(f) \\ &= \mathsf{V}[\widehat{R}_x(f)] + (\mathsf{E}[\widehat{R}_x(f)] - R_x(f))^2.\end{aligned}$$

9:3. From the periodogram, we can conclude that there seems to be two closely spaced sinusoids around $f = 0.3 - 0.35$ and from the modified periodogram we see a weaker sinusoid at $f = 0.1$. Our conclusion could then be that there are three sinusoids in the signal, two closely spaced and one weaker. However, we could not be sure if there are other signals that are even weaker or several more closely spaced around $f = 0.3$. The exercise illustrates the importance of using several estimation methods when dealing with unknown data. The true signal was

$$x(n) = \sin(2\pi 0.3n) + 0.5\sin(2\pi 0.33n) + 0.01\sin(2\pi 0.1n)$$

for $0 \leq n \leq 63$.

9:4. (a) For simplification we can restrict to $f = 0$ as the spectral density is constant $R_X(f) = \sigma^2$. For all other values of f, the result will then be the same. We get

$$\begin{aligned}\mathsf{E}[\widehat{R}_w(0)] &= \int_{-1/2}^{1/2} R_X(u)|W(u)|^2\,\mathrm{d}u \\ &= \sigma^2 \int_{-1/2}^{1/2} |W(u)|^2\,\mathrm{d}u = \sigma^2 \sum_{t=0}^{n-1} w^2(t),\end{aligned}$$

using Parseval's theorem. An unbiased spectral estimate is given if $\mathsf{E}[\widehat{R}_w(0)] = R_X(0) = \sigma^2$, i.e.,

$$\int_{-1/2}^{1/2} |W(u)|^2\,\mathrm{d}u = \sum_{t=0}^{n-1} w^2(t) = 1.$$

The usual periodogram without any data window, is unbiased for white noise.

(b) All covariances except the actual variance, are known to be zero for a white noise process, and to estimate $\hat{r}_x(0) = \hat{\sigma}^2$ we rely on that the integral of the spectral density gives the variance as,

$$r_X(0) = \int_{-1/2}^{1/2} R_X(f)df.$$

A periodogram is the estimated spectral density for discrete frequency values and if the periodogram is applied without zero-padding, the spectrum estimates for $f = k/n$, $k = 0,\ldots n/2-1$, are independent. We could note that the spectrum estimates for $k = n/2,\ldots,n-1$ are a mirrored copy for real-valued signals. Averaging the $n/2$ independent spectrum estimates,

$$\hat{\sigma}^2 = \frac{2}{n} \sum_{k=0}^{n/2-1} \hat{R}_x(k/n),$$

is an unbiased estimate of the white noise variance. As the variance of the periodogram is $V[\hat{R}_x(f)] \approx R_X^2(f) = \sigma^4$, the variance of the average will be

$$V[\hat{\sigma}^2] = V[\frac{2}{n} \sum_{k=0}^{n/2-1} \hat{R}_x(k/n)] \approx 2\sigma^4/n.$$

9:5. The answer is not easy since the Welch method gives considerable bias of the peak where the modified periodogram instead gives severe variance, which is seen as a large variation of the different estimates.

2018

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lu.se/