

Assignment 2, due September 22

Corrections (check the class message board): (none yet.)

1. Suppose that $Z_n \sim \mathcal{N}(0, 1)$ and

$$X_{n+1} = \frac{1}{2}X_n - \frac{5}{16}X_{n-1} + Z_n .$$

Show that this Gaussian process is stable. Find the limiting joint distribution of X_n and X_{n-1} in the limit $n \rightarrow \infty$.

2. (*Covariance of the covariance*) Suppose $X \sim \mathcal{N}(0, C)$ is a d component normal with mean zero and covariance C . Suppose we have independent samples $X_k \in \mathbb{R}^d$ and wish to estimate C . We discuss estimating $C = E[XX^t]$ using

$$\widehat{C} = \frac{1}{N} \sum_{k=1}^N X_k X_k^t . \quad (1)$$

Of course, \widehat{C} is a random variable with its own mean and covariance. The CLT implies that for large N , \widehat{C} is approximately a multivariate normal.

- (a) Show that the mean is $E[\widehat{C}] = C$.
- (b) (*Wick's formula for degree 4*) Suppose $X \sim \mathcal{N}(0, C)$ with components X_1, \dots, X_d (note the change of notation). Show that

$$E[X_i X_j X_k X_l] = C_{ij} C_{kl} + C_{ik} C_{jl} + C_{il} C_{jk} . \quad (2)$$

We saw the one dimensional version of this in the previous assignment. The method there applies here too. Another approach is to write a formula

$$x_l e^{x^t H x / 2} = - \sum_{m=1}^d A_{lm} \partial_{x_m} e^{-x^t H x / 2} .$$

The matrix A is related to $H = C^{-1}$ in some way. Now integrate by parts in the integral

$$\begin{aligned} E[X_i X_j X_k X_l] &= \frac{1}{Z} \int_{\mathbb{R}^d} x_i x_j x_k x_l e^{-x^t H x / 2} dx \\ &= \frac{1}{Z} \sum_{m=1}^d A_{lm} \int_{\mathbb{R}^d} \partial_{x_m} (x_i x_j x_k) e^{-x^t H x / 2} dx . \end{aligned}$$

(c) Find a formula for

$$D_{ij,kl} = \text{cov}(\widehat{C}_{ij}, \widehat{C}_{kl}) .$$

3. (*Maximum likelihood and the covariance*) Suppose $u(x, \theta)$ is a probability density that depends on an unknown parameter set θ . Suppose we have X_k , which are independent samples of u . The *maximum likelihood* estimate of θ is

$$\tilde{\theta} = \arg \max \prod_{k=1}^N u(X_k, \theta) . \quad (3)$$

Suppose $X \sim \mathcal{N}(0, C)$ and we want to estimate $H = C^{-1}$ from N independent samples. Show that the maximum likelihood estimator of C is given by the *empirical covariance* formula (1).

4. (*Kalman filter*) In the terminology of linear Gaussian processes, suppose

$$\begin{aligned} X_{n+1} &= AX_n + U_n && \text{(state dynamics)} \\ Y_n &= BX_n + V_n && \text{(observation)} \\ R &= \text{cov}(U_n) && \text{i.i.d., Gaussian} \\ S &= \text{cov}(V_n) && \text{i.i.d., Gaussian} \end{aligned}$$

Suppose $X_0 = 0$. Let C_n be the conditional variance of X_n , conditioned on knowing $Y_{1:n}$. Let \widehat{X}_n be the linear function of the observations $Y_{1:n}$ so that the prediction residual $X_n = \widehat{X}_n + W_n$ is independent of $Y_{1:n}$.

- (a) Show that $X_n = \widehat{X}_n + W_n$, where W_n is multivariate normal with mean 0 and covariance C_n , and independent of $Y_{1:n}$.
 (b) Find a matrix equation to solve for *gain* matrix K_{n+1} so that

$$\widehat{X}_{n+1} = A\widehat{X}_n + K_{n+1} (Y_{n+1} - \widehat{Y}_{n+1}) .$$

where \widehat{Y}_{n+1} is the predicted observation based on $Y_{1:n}$, given by $\widehat{Y}_{n+1} = BA\widehat{X}_n$. The equation will involve the five relevant matrices: A , B , R , S , and C_n . Find a formula for C_{n+1} . Hint, once the prediction residual at time $n + 1$ is uncorrelated to $Y_{n+1} - \widehat{Y}_{n+1}$, you have it, why?

- (c) Assume $C_n \rightarrow C$ as $n \rightarrow \infty$ and $K_n \rightarrow K$ as $n \rightarrow \infty$. Find nonlinear equations for C and K . These are called *matrix Riccati equations*. Control theory is full of them.
 (d) Download the code `GaussianProcess.R` and see what it does. Modify it to do Kalman filtering to make the best estimate of X_n measuring only $X_{1:n}$ with measurement error variance σ^2 . Compare the long time predicted mean square prediction error with the theoretical prediction of what the mean square prediction error should be.