1 Jointly Gaussian

Problem 1. Prove (d) of Note 5.

Solution. We note that \( Y - L[Y|Z] \perp Z \), and \( E[Y - L[Y|Z]] = 0 \) hence
\[
\]
It follows that \( Y - L[Y|Z] \) and \( Z \) are uncorrelated. Since they are both linear transformations of the Gaussian r.v. \( X = (Y, Z) \), then \( Y - L[Y|Z] \) and \( Z \) are jointly Gaussian. This implies that they are also independent, hence \( Y - L[Y|Z] \) and \( \phi(Z) \) are independent for all functions \( \phi(\cdot) \). Then \( Y - L[Y|Z] \) and \( \phi(Z) \) are uncorrelated for all functions \( \phi(\cdot) \). Since \( Y - L[Y|Z] \) is an unbiased estimator we have
\[
E[(Y - L[Y|Z])\phi(Z)] = \text{cov}(Y - L[Y|Z], \phi(Z)) + E[Y - L[Y|Z]]E[\phi(Z)] = 0
\]
hence \( Y - L[Y|Z] \perp \phi(Z), \ \forall \phi(\cdot), \) so \( L[Y|Z] = E[Y|Z] \).

Problem 2. Let \( X \) be the height of the father, \( Y \) the height of the son, in a sample of father-son pairs. Assume \( X \) and \( Y \) bivariate normal, as found by Karl Pearson around 1900. Assume \( E(X) = 68 \) (inches), \( E(Y) = 69 \), \( \sigma_X = \sigma_Y = 2 \), \( \rho = .5 \). (We expect \( \rho \) to be positive because on the average, the taller the father, the taller the son.) Given \( X = 80 \) (6 feet 8 inches), what is the distribution of \( Y \)?

Solution. \( Y \) is normal with mean
\[
\mu_Y + \frac{\rho \sigma_Y}{\sigma_X}(x \mu_X) = 69 + .5(8068) = 75
\]
which is 6 feet 3 inches. The variance of \( Y \) given \( X = 80 \) is
\[
\sigma_Y^2(1^2) = 4(3/4) = 3.
\]
Thus the son will tend to be of above average height, but not as tall as the father. This phenomenon is often called regression, and the line \( y = \mu_Y + \frac{\rho \sigma_Y}{\sigma_X}(x \mu_X) \) is called the regression line.

Problem 3. Let \( X, Y, Z \) be i.i.d. \( \mathcal{N}(0,1) \). Find
\[
E[X \mid X + Y, X + Z, Y - Z]
\]
Hint: Argue that the observation \( Y - Z \) is redundant.
**Solution.** The first thing to notice is that \( Y - Z = (X+Y) - (X+Z) \), so that the three observations are linearly dependent. Thus,

\[
E[X|X+Y, X+Z, Y-Z] = E[X|X+Y, X+Z].
\]

We could use the standard approach and compute the covariances. Instead, we note that, by symmetry,

\[
E[X|X+Y, X+Z] = a(X+Y) + a(X+Z).
\]

The coefficient \( a \) is such that the error is orthogonal to the observations. Thus,

\[
E((X - a(X+Y) - a(Y+Z))(X+Y)) = 0.
\]

This gives

\[
1 - 2a - a = 0,
\]

so that \( a = 1/3 \) and we have

\[
E[X|X+Y, X+Z, Y-Z] = \frac{1}{3}(X+Y) + \frac{1}{3}(X+Z).
\]

## 2 Kalman filter

**Problem 4.** Formulate the falling object problem as a Kalman filter.

**Solution.** Let \( z(t) \) be the altitude of the object, then from Newton’s 2nd law of motion

\[
m \frac{\partial^2 z}{\partial t^2} = F = -mg
\]

because the object is only subject to gravity. Integrating this equation two times gives:

\[
z(t) = -\frac{gt^2}{2} + s_0t + z_0, \quad t \geq 0
\]

where \( s_0 \) and \( z_0 \) are the initial velocity and altitude of the object respectively. Then the discretized dynamics of the system are

\[
z_n = -\frac{gn^2}{2} + s_0n + z_0, \quad n \geq 0
\]

We observe

\[
\eta_n = z_n + w_n
\]

where \( w_n \) is some noise with variance \( r_n \). Note that \( s_0 \) and \( z_0 \) are the controls of the system. Since the term \( -\frac{gn^2}{2} \) is known, we do the change of variables \( x_n = z_n + \frac{gn^2}{2} \) and \( y_n = \eta_n + \frac{gn^2}{2} \) and the equations become \( x_{n+1} = x_n + s_0 \) and \( y_n = x_n + w_n \) where \( v_n \) is a zero-mean process noise with variance \( q_k \) that covers unmodelled dynamics (e.g. drag).

**Predict**

- Predicted (a priori) state estimate \( x_{n+1|n} = x_{n|n} + s_0 \)
- Predicted (a priori) estimate variance \( \sigma_{n+1|n} = \sigma_{n|n} + q_n \)
Update

Innovation or measurement residual:
\[ \epsilon_{n+1} = y_{n+1} - x_{n+1|n} \]

Innovation (or residual) covariance:
\[ u_{n+1} = \sigma_{n+1|n} + r_{n+1} \]
\[ k_{n+1} = \frac{\sigma_{n+1|n}}{u_{n+1}} = \frac{\sigma_{n+1|n}}{(\sigma_{n+1|n} + r_{n+1})} \]
\[ x_{n+1|n+1} = x_{n+1|n} + k_{n+1}\epsilon_{n+1} = (1 - k_{n+1})x_{n+1|n} + k_{n+1}y_{n+1} \]
\[ \sigma_{n+1|n+1} = (1 - k_{n+1})\sigma_{n+1|n} \]

Optional Kalman gain

Updated (a posteriori) state estimate:
\[ x_{n+1|n+1} = x_{n+1|n} + k_{n+1}y_{n+1} \]

Updated (a posteriori) estimate covariance:
\[ \sigma_{n+1|n+1} = (1 - k_{n+1})\sigma_{n+1|n} \]

We model the trajectory of the falling object as a parabola, to which we add a drag due to air friction proportional to the velocity of the object. We suppose that the dynamics associated to air resistance are unknown, so we cover them with a Gaussian noise with variance proportional to the velocity of the object.

![Figure 1: Prediction without observations. Estimate when including observations.](image)

We notice that the prediction model fails to give a good estimate of the trajectory of the falling object because it does not include unmodelled dynamics (here, air friction). However, the Kalman filter performs well despite the noisy measurements.

![Figure 2: Kalman gain and variance of the unmodelled dynamics.](image)

We notice that when the variance of the unmodelled dynamics is high, the Kalman gain is also high, i.e. we put more trust in the observations.

**Problem 5.** In a linear system with independent Gaussian noise, with state \( X_n \) and observation
$Y_n$, the Kalman filter computes (choose the correct answers, if any)

□ MLE[$Y_n|X_n$];    □ MLE[$X_n|Y_n$];    □ MAP[$Y_n|X_n$];    □ MAP[$X_n|Y_n$];

□ E[$X_n|Y_n$];    □ E[$Y_n|X_n$];    □ E[$X_n|Y_n$];    □ E[$Y_n|X_n$];

**Solution.** The Kalman filter computes $\hat{X}_n = E[X_n|Y_n]$. This happens to be also $MAP[X_n|Y_n]$ because, given $Y_n$, the random variable $X_n$ is Gaussian with mean $\hat{X}_n$, so that its most likely value is the mean $\hat{X}_n$. The other possibilities are not correct.