1. **Update for vector Kalman filter** Assume that \( X \) and \( Y \) are zero-mean, jointly Gaussian, jointly non-singular, related by \( Y = HX + Z \). Let \( \hat{x}(y) \) be the MMSE estimator and let \( \xi = X - \hat{x}(y) \).

(a) Show that
\[
\hat{x}(y) = K_\xi H^T K_Z^{-1} y.
\]

(b) Show that
\[
K_\xi^{-1} = K_X^{-1} + H^T K_Z^{-1} H.
\]

(Hint: you can use the following dual representation: for jointly Gaussian random vectors \( X \) and \( Y \) one can write \( Y = HX + Z \) and \( X = GY + V \), simultaneously. Then you can use the fact that \( K_Y^{-1}G = H^T K_Z^{-1} \).)

**Solution:** Part (a)

\[
\text{Solution: As shown in Section 3.5, the relation between } X \text{ and } Y \text{ can be expressed both as } Y = [H]X + Z \text{ and } X = [G]Y + V \text{ with the symmetric relations}
\]
\[
[G] = [K_{X,Y} K_Y^{-1}]; \quad [H] = [K_{X,Y} K_X^{-1}]. \tag{A.225}
\]

From the relationship \( X = [G]y + V \), it is clear that the MMSE estimate of \( X \) for a sample observation \( y \) is \( \hat{x}(y) = [G]y \) and that the estimation error is \( v \) and thus that \( [K_\xi] = [K_V] \).

The matrices \([G]\) and \([H]\) are related in (3.115) as \([G^T K_V^{-1}] = [K_Z^{-1} H] \) and thus
\[
[G^T] = [K_Z^{-1} HK_V] = [K_Z^{-1} HK_\xi]. \tag{A.226}
\]

Taking the transpose of this, we get (A.224).

Part (b)

**Solution:** Following the hint,
\[
[K_{X}^{-1}] = [K_\xi^{-1}] - [K_X^{-1}]K_{X,Y} K_Y^{-1} K_{X,Y}^{-1} [K_{X}^{-1}].
\]

From (A.225), we can substitute \([H^T]\) for \([K_X^{-1}]K_{X,Y}\) and \([G^T]\) for \([K_Y^{-1}]K_{X,Y}\) above, getting
\[
[K_{X}^{-1}] = [K_\xi^{-1}] - [H^T G^T K_\xi^{-1}]. \tag{A.227}
\]

Rearranging and substituting (A.226) into this, we get (A.227).

One minor detail we have omitted above is to verify that \([K_Z]\) and \([K_V]\) are non-singular. If \([K_Z]\) is singular, then there is some vector \( a \) such that \( a^T Z = 0 \). This implies that \( a^T X = [G] Y \), which means that \( X \) and \( Y \) can not be mutually non-singular. Since we have assumed that \( X \) and \( Y \) are mutually non-singular, this means that \( Z \) (and thus \([K_Z]\)) is non-singular. The same argument applies to \( V \).

2. **Kalman filter for location tracking** Consider a truck on frictionless, straight rails \(^1\). Initially,

\(^1\)Example taken from https://en.wikipedia.org/wiki/Kalman_filter#Example_application
the truck is stationary at location $L_0 = 0$, but it is buffeted by random uncontrolled forces. We measure the position of the truck every $\Delta t = 1$ seconds, but these measurements are imprecise; we want to maintain a model of the truck’s location $L_t$ and its velocity $V_t$.

Specifically, assuming at time $t = 0$, the initial state of the truck is $L_0 = 0$, $V_0 = 1$. Between $t − 1$ and $t$, the velocity is subject to a constant acceleration $A_{t−1} \sim \mathcal{N}(0, \sigma_a^2)$. Also at time $t$, we take a noisy observation $Y_t = L_t + Z_t$, where $Z_t \sim \mathcal{N}(0, \sigma_z^2)$. Formulate this problem as a vector Kalman filter estimation problem.

**Solution:** According to Newton’s law, the dynamic can be written as

\[
L_t = L_{t-1} + V_{t-1} + \frac{1}{2}A_{t-1}
\]

\[
V_t = V_{t-1} + A_{t-1}
\]

The observation can be written as

\[
Y_t = L_t + Z_t.
\]

Define the state at time $t$, $X_t = [L_t, V_t]^T$. Then the above can be written as

\[
X_t = \begin{bmatrix} L_t \\ V_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X_{t-1} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} A_{t-1}
\]

\[
Y_t = [1, 0]X_t + Z_t.
\]

3. **LLSE and linear regression.** Consider a set of $n$ pairs of numbers $\{(x_i, y_i), i = 1, \ldots, n\}$. Linear regression is the (non-probabiltistic) problem of finding a line $z_i = \alpha x_i + \beta$ through the points that minimizes the mean square error (MSE)

\[
\epsilon_{\text{MSE}} = \sum_{i=1}^{n} (z_i - y_i)^2
\]

(a) Find the values of $\alpha$ and $\beta$ that minimize MSE by differentiating $\epsilon_{\text{MSE}}$.

(b) Show that if $(x_i, y_i)$ are realizations of a pair of random variables $(X, Y)$, $z_i$ approaches the LLSE estimator for $Y|X = x_i$ in the limit of large $n$.

**Solution:**

(a) We find that

\[
\frac{d\epsilon_{\text{MSE}}}{d\alpha} = 2 \sum_{i=1}^{n} x_i (\alpha x_i + \beta - y_i)
\]

\[
\frac{d\epsilon_{\text{MSE}}}{d\beta} = \sum_{i=1}^{n} (\alpha x_i + \beta - y_i)
\]

which can be readily set to zero and solved. We have

\[
\beta = \frac{1}{n} \sum_{i=1}^{n} y_i - \frac{1}{n} \sum_{i=1}^{n} x_i, \quad \alpha = \frac{\frac{1}{n} \sum_{i=1}^{n} x_i y_i - (\frac{1}{n} \sum_{i=1}^{n} x_i) (\frac{1}{n} \sum_{i=1}^{n} y_i)}{\frac{1}{n} \sum_{i=1}^{n} x_i^2 - (\frac{1}{n} \sum_{i=1}^{n} x_i)^2}
\]

Each $z_m$ approximates $y_m$ as

\[
z_m = \frac{1}{n} \sum_{i=1}^{n} y_i + \frac{\frac{1}{n} \sum_{i=1}^{n} x_i y_i - (\frac{1}{n} \sum_{i=1}^{n} x_i) (\frac{1}{n} \sum_{i=1}^{n} y_i)}{\frac{1}{n} \sum_{i=1}^{n} x_i^2 - (\frac{1}{n} \sum_{i=1}^{n} x_i)^2} \left(x_m - \frac{1}{n} \sum_{i=1}^{n} x_i\right)
\]
(b) If we regard each \((x_i, y_i)\) as realizations of a pair of random variables, then by the law of large numbers

\[
z_m \to \mathbb{E}[Y] + \frac{K_{XY}}{\sigma^2_X} (x_m - \mathbb{E}[X]) = \text{LLSE}[Y|X = x_m]
\]

in the limit of large \(n\).