of enlarging the number \( m \) of observations, and essentially shows that the estimation error must approach 0 as \( m \to \infty \).

### 10.10 Exercises

**Exercise 10.1** (a) Consider the joint probability density \( f_{X,Z}(x,z) = e^{-z} \) for \( 0 \leq x \leq z \) and \( f_{X,Z}(x,z) = 0 \) otherwise. Find the pair \( x,z \) of values that maximize this density. Find the marginal density \( f_Z(z) \) and find the value of \( z \) that maximizes this.

(b) Let \( f_{X,Y}(x,z,y) = y^2 e^{-yz} \) for \( 0 \leq x \leq z \), \( 1 \leq y \leq 2 \) and be 0 otherwise. Conditional on an observation \( Y = y \), find the joint MAP estimate of \( X,Z \). Find the marginal density \( f_{Z|Y}(z|y) \), the marginal density of \( Z \) conditional on \( Y = y \), and find the MAP estimate of \( Z \) conditional on \( Y = y \).

**Exercise 10.2** Let \( Y = X + Z \), where \( X \) and \( Z \) are IID and each \( \mathcal{N}(0,1) \). Let \( U = Z - X \).

(a) Explain (using the results of Chapter 3) why \( Y \) and \( U \) are jointly Gaussian and why they are statistically independent.

(b) Without using any matrices, write down the joint probability density of \( Y \) and \( U \). Verify your answer from (3.22).

(c) Find the MMSE estimate \( \hat{x}(y) \) of \( X \) conditional on a given sample value \( y \) of \( Y \). You can derive this from first principles, or use (10.9) or Example 10.2.2.

(d) Show that the estimation error \( \xi = X - \hat{X}(Y) \) is equal to \(-U/2\).

(e) Write down the probability density of \( U \) conditional on \( Y = y \) and that of \( \xi \) conditional on \( X = y \).

(f) Draw a sketch in the \( x,z \) plane of the equiprobability contours of \( X \) and \( Z \). Explain why these are also equiprobability contours for \( Y \) and \( U \). For some given sample value of \( Y \), say \( Y = 1 \), illustrate the set of points for which \( x + z = 1 \). For a given point on this line, illustrate the sample value of \( U \).

**Exercise 10.3** (a) Let \( X,Z_1,Z_2,\ldots,Z_n \) be independent zero-mean Gaussian \( \text{rvs} \) with variances \( \sigma_X^2,\sigma_{Z_1}^2,\ldots,\sigma_{Z_n}^2 \) respectively. Let \( Y_j = h_jX + Z_j \) for \( j \geq 1 \) and let \( Y = (Y_1,\ldots,Y_n)^T \). Use (10.9) to show that the MMSE estimate of \( X \) conditional on \( Y = y = (y_1,\ldots,y_n)^T \) is given by

\[
\hat{x}(y) = \sum_{j=1}^n g_j y_j, \quad \text{where} \quad g_j = \frac{h_j/\sigma_{Z_j}^2}{(1/\sigma_X^2) + \sum_{i=1}^n h_i^2/\sigma_{Z_i}^2}.
\]  

(10.115)

Hint: Let the row vector \( g^i \) be \( [K_{XY}][K_Y^{-1}] \) and multiply \( g^i \) by \( K_Y \) to solve for \( g^i \).

(b) Let \( \xi = X - \hat{X}(Y) \) and show that (10.29) is valid, i.e., that

\[
1/\sigma_\xi^2 = 1/\sigma_X^2 + \sum_{i=1}^n \frac{h_i^2}{\sigma_{Z_i}^2}.
\]

(c) Show that (10.28), i.e., \( \hat{x}(y) = \sigma_X^2 \sum_{j=1}^n h_j y_j/\sigma_{Z_j}^2 \), is valid.

(d) Show that the expression in (10.29) is equivalent to the iterative expression in (10.27).
(e) Show that the expression in (10.28) is equivalent to the iterative expression in (10.25).

Exercise 10.4 Let \( X, Z_1, \ldots, Z_n \) be independent, zero-mean, and Gaussian, satisfying \( Y_i = h_iX + Z_i \) for \( 1 \leq i \leq n \). The point of this exercise is to show that estimation of \( X \) from \( Y^n_1 \) is essentially no more complex than the same problem with each \( h_i = 1 \).

(a) Write out the MMSE estimate and variance of the estimation error from (10.28) and (10.29) where each \( h_i = 1 \).

(b) For arbitrary numbers \( h_1, \ldots, h_n \), consider the equations \( h_iY_i = h_iX + h_iZ_i \). Let \( U_i = h_iY_i \). Write out the MMSE estimate of \( X \) as a function of \( U_1, \ldots, U_n \), still using the \( n \) \( Z_i \)'s. Hint: Note that if you scale each observation \( U_i \) to \( U_i/h_i \) you have the same estimation problem as in (a).

(c) Now let \( W_i = h_iZ_i \). Write out the same MMSE estimate in terms of both \( U_i, \ldots, U_n \) and \( W_1, \ldots, W_n \) and show that you now have (10.28) with all the \( h_i \)'s back in the equation. Hint: You still have the basic equation in (a), but the variance of the \( Z_i \)'s must be expressed in terms of the \( W_i \)'s. Write out the error variance from (10.29) in the same way.

Exercise 10.5 (a) Assume that \( X_1 \sim \mathcal{N}(\overline{X}, \sigma^2_{X_1}) \) and that for each \( n \geq 1 \), \( X_{n+1} = \alpha X_n + W_n \), where \( 0 < \alpha < 1 \), \( W_n \sim \mathcal{N}(0, \sigma^2_W) \), and \( X_1, W_1, W_2, \ldots, \) are independent. Show that for each \( n \geq 1 \),

\[
E[X_n] = \alpha^{n-1} \overline{X}, \quad \sigma^2_{X_n} = \frac{(1 - \alpha^{2(n-1)})\sigma^2_W}{1 - \alpha^2} + \alpha^{2(n-1)}\sigma^2_X.
\]

(b) Show directly, by comparing the equation \( \sigma^2_{X_n} = \alpha^2\sigma^2_{X_{n-1}} + \sigma^2_W \) for each two adjoining values of \( n \), that \( \sigma^2_{X_n} \) moves monotonically from \( \sigma^2_{X_1} \) to \( \sigma^2_W/(1-\alpha^2) \) as \( n \to \infty \).

(c) Assume that sample values of \( Y_1, Y_2, \ldots, \) are observed, where \( Y_n = hX_n + Z_n \) and where \( Z_1, Z_2, \ldots, \) are IID zero-mean Gaussian \( n \) \( Z \)'s with variance \( \sigma^2_Z \). Assume that \( Z_1, \ldots, W_1, \ldots, X_1 \) are independent and assume \( h \geq 0 \). Rewrite the recursion for the variance of the estimation error in (10.41) for this special case. Show that if \( h/\sigma_Z = 0 \), then \( \sigma^2_{\xi_n} = \sigma^2_{\xi_n} \) for each \( n \geq 1 \). Hint: Compare the recursion in (b) to that for \( \sigma^2_{\xi_n} \).

(d) Show from the recursion that \( \sigma^2_{\xi_n} \) is a decreasing function of \( h/\sigma_Z \) for each \( n \geq 2 \). Use this to show that \( \sigma^2_{\xi_n} \leq \sigma^2_{\xi_1} \) for each \( n \). Explain (without equations) why this result must be true.

(e) Show that the sequence \( \{\sigma^2_{\xi_n}, n \geq 1\} \) is monotonic in \( n \). Hint: Use the same technique as in (b). From this and (d), show that \( \lambda = \lim_{n \to \infty} \sigma^2_{\xi_n} \) exists. Show that the limit satisfies (10.42) (note that (10.42) must have two roots, one positive and one negative, so the limit must be the positive root).

(f) Show that for each \( n \geq 1 \), \( \sigma^2_{\xi_n} \) is increasing in \( \sigma^2_W \) and increasing in \( \alpha \). Note: This increase with \( \alpha \) is surprising, since when \( \alpha \) is close to 1, \( X_n \) changes slowly, so we would expect to be able to track \( X_n \) well. The problem is that \( \lim_n \sigma^2_{X_n} = \sigma^2_W/(1 - \alpha^2) \) so the variance of the untracked \( X_n \) is increasing without limit as \( \alpha \) approaches 1. Part (g) is somewhat messy, but resolves this issue.

(g) Show that if the recursion is expressed in terms of \( \beta = \sigma^2_W/(1 - \alpha^2) \) and \( \alpha \), then \( \lambda \) is decreasing in \( \alpha \) for constant \( \beta \).
Exercise 10.6  (a) Assume that $X$ and $Y$ are zero-mean, jointly Gaussian, jointly non-singular, and related by $Y = [H]X + Z$. Show that the MMSE estimate can be expressed as

$$
\hat{x}(y) = [K_y H'] [K_Z^{-1}] y.
$$

Hint: Recall from (10.9) that $\hat{x}(y) = [G] y$. Relate $[G]$ and $[H]$ by using (3.115).

(b) Show that

$$
[K_y^{-1}] = [K_X^{-1}] + [H' K_Z^{-1} H].
$$

Hint: A reasonable approach is to start with (10.10), i.e., $[K_y] = [K_X] - [K_X Y K_Y^{-1} K_X Y]$. Premultiply both sides by $[K_X^{-1}]$ and post-multiply both sides by $[K_y^{-1}]$. The final product of matrices here can be expressed as $[H' K_Z^{-1} H]$ by using the hint in (a) on $[K_Z^{-1} H]$.

Exercise 10.7  (a) Write out $E \left[(X - g'Y)^2\right] = \sigma_X^2 - 2[K_{XY}] g + g'[K_Y] g$ as a function of $g = (g_1, g_2, \ldots, g_n)'$ and take the partial derivative of the function with respect to $g_i$ for each $i$, $1 \leq i \leq n$. Show that the vector of these partial derivatives is $-2[K_{XY}] + 2 a'[K_Y]$.

(b) Explain why the stationary point here is actually a minimum.

Exercise 10.8  For a real inner-product space, show that $m$ vectors, $Y_1, \ldots, Y_m$ are linearly dependent if and only if the matrix of inner products, $\{(Y_j, Y_k); 1 \leq j, k \leq m\}$, is singular.

Exercise 10.9  Show that (10.89) to (10.91) agree with (10.11).

Exercise 10.10  Let $(X = X_1, \ldots, X_n)'$ be a zero-mean complex rv with real and imaginary components $X_{re,j}, X_{im,j}$, $1 \leq j \leq n$ respectively. Express $E[X_{re,j}X_{re,k}], E[X_{re,j}X_{im,k}], E[X_{im,j}X_{im,k}], E[X_{im,j}X_{re,k}]$ as functions of the components of $[K_X]$ and $E[XX']$.

Exercise 10.11  Let $Y = Y_r + j Y_{im}$ be a complex rv. For arbitrary real numbers $a, b, c, d$, find complex numbers $\alpha$ and $\beta$ such that

$$
\Re \left[\alpha Y + \beta Y^*\right] = a Y_{re} + b Y_{im},
$$

$$
\Im \left[\alpha Y + \beta Y^*\right] = c Y_{re} + d Y_{im}.
$$

Exercise 10.12 (Derivation of circularly-symmetric Gaussian density)  Let $X = X_{re} + i X_{im}$ be a zero-mean, circularly-symmetric, $n$-dimensional, Gaussian complex rv. Let $U = (X_{re}', X_{im}')'$ be the corresponding $2n$-dimensional real rv. Let $[K_{re}] = E [X_{re} X'_{re}]$ and $[K_{ri}] = E [X_{re} X'_{im}]$.

(a) Show that

$$
[K_U] = \begin{bmatrix}
[K_{re}] & [K_{ri}] \\
-[K_{ri}] & [K_{re}]
\end{bmatrix}.
$$

(b) Show that

$$
[K_U^{-1}] = \begin{bmatrix}
B & C \\
-C & B
\end{bmatrix}
$$

and find the $B, C$ for which this is true.
(c) Show that \( [K_X] = 2([K_{re}] - [K_{ri}] ) \).

(d) Show that \( [K_X^{-1}] = \frac{1}{2}(B - iC) \).

(e) Define \( f_X(x) = f_U(u) \) for \( u = (x_{re}, x_{im}) \) and show that
\[
f_X(x) = \frac{\exp\left(-x^\dagger [K_X^{-1}] x^\dagger\right)}{(2\pi)^n \sqrt{\det[K_U]}}.
\]

(f) Show that
\[
\det[K_U] = \begin{bmatrix}
[K_{re}] + i[K_{ri}] & [K_{ri}] - i[K_{re}] \\
-[K_{ri}] & [K_{re}]
\end{bmatrix}.
\]

Hint: Recall that elementary row operations do not change the value of a determinant.

(g) Show that
\[
[K_U] = \begin{bmatrix}
[K_{re}] + i[K_{ri}] & 0 \\
-[K_{ri}] & [K_{re}]
\end{bmatrix}.
\]

Hint: Recall that elementary column operations do not change the value of a determinant.

(h) Show that
\[
\det[K_U] = 2^{-2n} (\det[K_X])^2
\]
and from this conclude that (3.108) is valid.

Exercise 10.13 (Alternate derivation of circularly-symmetric Gaussian density)

(a) Let \( X \) be a circularly-symmetric, zero-mean, complex Gaussian rv with covariance 1. Show that
\[
f_X(x) = \frac{\exp(-x^\dagger x)}{\pi}.
\]
Recall that the real part and imaginary part each have variance 1/2.

(b) Let \( X \) be an \( n \)-dimensional, circularly-symmetric, complex Gaussian zero-mean rv with \( [K_X] = I_n \). Show that
\[
f_X(x) = \frac{\exp(-x^\dagger x)}{\pi^n}.
\]

(c) Let \( Y = [H]X \) where \( [H] \) is \( n \times n \) and invertible. Show that
\[
f_Y(y) = \frac{\exp\left[-y^\dagger ([H^{-1}]^\dagger [H^{-1}] y\right]}{v\pi^n},
\]
where \( v \) is \( dy/dx \), the ratio of an incremental \( 2n \)-dimensional volume element after being transformed by \( [H] \) to that before being transformed.

(d) Use this to show that that (3.108) is valid.

Exercise 10.14 (a) Let \( Y = X^2 + Z \), where \( Z \) is a zero-mean unit variance Gaussian rv. Show that no unbiased estimate of \( X \) exists from observation of \( Y \). Hint. Consider any \( x > 0 \) and compare with \(-x\).
(b) Let \( Y = X + Z \), where \( Z \) is uniform over \((-1, 1)\) and \( X \) is a parameter lying in \((-1, 1)\). Show that \( \hat{x}(y) = y \) is an unbiased estimate of \( x \). Find a biased estimate \( \hat{x}_1(y) \) for which \( |\hat{x}_1(y) - x| \leq |\hat{x}(y) - x| \) for all \( x \) and \( y \) with strict inequality with positive probability for all \( x \in (-1, 1) \).

Exercise 10.15  
(a) Assume that for each parameter value \( x \), \( Y \) is Gaussian, \( \mathcal{N}(x, \sigma_Z^2) \). Show that \( V_x(y) \) as defined in \( (10.103) \) is equal to \( (y - x)/\sigma_Z^2 \) and show that the Fisher information is equal to \( 1/\sigma_Z^2 \).

(b) Show that for ML estimation, the bias is 0 for all \( x \). Show that the Cramer–Rao bound is satisfied with equality for all \( x \) in this case.

(c) Consider using the MMSE estimator for the a priori distribution \( X \sim \mathcal{N}(0, \sigma_X^2) \). Show that the bias satisfies \( b_z(x) = -x\sigma_Z^2/(\sigma_Z^2 + \sigma_X^2) \).

(d) Show that the MMSE estimator in (c) satisfies the Cramer–Rao bound with equality for each \( x \). Note that the mean-squared error, as a function of \( x \), is smaller than that for the ML case for small \( x \) and larger for large \( x \).

Exercise 10.16  
Assume that \( Y \) is \( \mathcal{N}(0, x) \). Show that \( V_x(y) \) as defined in \( (10.103) \) is \( V_x(y) = [y^2/x - 1]/(2x) \). Verify that \( V_x(Y) \) is zero mean for each \( x \). Find the Fisher information, \( J(x) \).