# STAT 665 - HW 1

*Maggie Buffum*

## Problem 5.1

Let $x = (X_1, \dots, X_k) \sim N_k(\mu, \Sigma)$, with $r(\Sigma) = k$.

(a) Show that

$$I = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} exp\Big\{ -\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\Big\} dx_1 \dots dx_k$$
$$= (2\pi)^{k/2}|\Sigma|^{1/2}.$$

---

What we have is part of a multivariate normal distribution integrated over all possible values of $\mathbf{x}$. Let's first note that by definition of positive-definite, $\Sigma^{-1}$ must be a positive definite matrix, otherwise its inverse would not exist. Now, we know that

$$f(\mathbf{x};\mu,\Sigma) = \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} exp\Big\{ -\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\Big\} \text{ for all } \mathbf{x} \in \mathcal{R}^k$$

and the integration of $f(\mathbf{x};\mu,\Sigma)$ over all possible values of $\mathbf{x}$ is 1. Therefore, we have that

$$1 = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} exp\Big\{ -\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\Big\} dx_1 \dots dx_k$$
$$\implies (2\pi)^{k/2}|\Sigma|^{1/2} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} exp\Big\{ -\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\Big\} dx_1 \dots dx_k$$

(b) Evaluate $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} exp\{-(x_1^2 + 2x_1x_2 + 4x_2^2)\} dx_1 dx_2$.

---

Let's rewrite the expression in the following way:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} exp\{-(x_1^2 + 2x_1x_2 + 4x_2^2)\} dx_1 dx_2$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} exp\{-\frac{1}{2}(2x_1^2 + 4x_1x_2 + 8x_2^2)\} dx_1 dx_2$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} exp\{-\frac{1}{2}(\mathbf{x}'\mathbf{A}\mathbf{x})\} dx_1 dx_2$$

where $\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 2 & 8 \end{bmatrix}$, a $2 \times 2$ positive-definite matrix. Applying Aitken's integral, we have the solution

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} exp\{-\frac{1}{2}(\mathbf{x}'\mathbf{A}\mathbf{x})\} dx_1 dx_2$$
$$= (2\pi)^{2/2}|\mathbf{A}|^{-1/2}$$
$$= (2\pi)\Bigg| \begin{bmatrix} 2 & 2 \\ 2 & 8 \end{bmatrix} \Bigg|^{-1/2}$$
$$= (2\pi)((2)(8) - (2)(2))^{-1/2}$$
$$= (2\pi)(12)^{-1/2}$$
$$= (2\pi)\frac{2}{\sqrt{3}}$$
$$= \frac{\pi}{\sqrt{3}}$$

## Problem 5.2

[Graybill, 1961]. Let $x = (X_1, X_2)$ have a bivariate normal distribution with pdf

$$f(x; \mu, \Sigma) = \frac{1}{k} exp[-Q/2]$$

where $Q = 2x_1^2 - x_1x_2 + 4x_2^2 - 11x_1 - 5x_2 + 19$, and $k$ is a constant. Find a constant $a$ such that $P(3X_1 - X_2 < a) = 0.01$.

---

Since we know the distribution of the vector $\mathbf{x}$ is normal, $Q = (\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu)$. We can solve for $\Sigma$ as

$$\begin{aligned}
Q &= \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & \sigma_{12} \\ \sigma_{12} & \sigma_2 \end{bmatrix}^{-1} \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \frac{1}{|\Sigma|} \begin{bmatrix} \sigma_2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1 \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \\
&= \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \begin{bmatrix} \frac{\sigma_2}{|\Sigma|} & -\frac{\sigma_{12}}{|\Sigma|} \\ -\frac{\sigma_{12}}{|\Sigma|} & \frac{\sigma_1}{|\Sigma|} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \\
&= \begin{bmatrix} (x_1 - \mu_1)\frac{\sigma_2}{|\Sigma|} - (x_2 - \mu_2)\frac{\sigma_{12}}{|\Sigma|} & -(x_1 - \mu_1)\frac{\sigma_{12}}{|\Sigma|} + (x_2 - \mu_2)\frac{\sigma_1}{|\Sigma|} \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \\
&= (x_1 - \mu_1)^2 \frac{\sigma_2}{|\Sigma|} - (x_2 - \mu_2)\frac{\sigma_{12}}{|\Sigma|}(x_1 - \mu_1) - (x_1 - \mu_1)\frac{\sigma_{12}}{|\Sigma|}(x_2 - \mu_2) + (x_2 - \mu_2)^2 \frac{\sigma_1}{|\Sigma|} \\
&= \frac{\sigma_2}{|\Sigma|}(x_1 - \mu_2)^2 + \frac{\sigma_1}{|\Sigma|}(x_2 - \mu_2)^2 - 2\frac{\sigma_{12}}{|\Sigma|}(x_1 - \mu_1)(x_2 - \mu_2) \\
&= \frac{\sigma_2}{|\Sigma|}(x_1^2 + \mu_1^2 - 2x_1\mu_1) + \frac{\sigma_1}{|\Sigma|}(x_2^2 + \mu_2^2 - 2x_2\mu_2) - 2\frac{\sigma_{12}}{|\Sigma|}(x_1x_2 - x_1\mu_2 - x_2\mu_1 + \mu_1\mu_2) \\
&= \frac{\sigma_2}{|\Sigma|}x_1^2 + \frac{\sigma_1}{|\Sigma|}x_2^2 - 2\frac{\sigma_{12}}{|\Sigma|}x_1x_2 + \frac{\sigma_2}{|\Sigma|}\mu_1^2 - 2\frac{\sigma_2}{|\Sigma|}x_1\mu_1 + \frac{\sigma_1}{|\Sigma|}\mu_2^2 - 2\frac{\sigma_1}{|\Sigma|}x_2\mu_2 + 2\frac{\sigma_{12}}{|\Sigma|}x_1\mu_2 + 2\frac{\sigma_{12}}{|\Sigma|}x_2\mu_1 - 2\frac{\sigma_{12}}{|\Sigma|}\mu_1\mu_2 \\
&= \frac{\sigma_2}{|\Sigma|}x_1^2 + \frac{\sigma_1}{|\Sigma|}x_2^2 - 2\frac{\sigma_{12}}{|\Sigma|}x_1x_2 + x_1\Bigg( - 2\frac{\sigma_2}{|\Sigma|}\mu_1 + 2\frac{\sigma_{12}}{|\Sigma|}\mu_2\Bigg) + x_2\Bigg( - 2\frac{\sigma_1}{|\Sigma|}\mu_2 + 2\frac{\sigma_{12}}{|\Sigma|}\mu_1\Bigg) + \frac{\sigma_2}{|\Sigma|}\mu_1^2 + \frac{\sigma_1}{|\Sigma|}\mu_2^2 - 2\frac{\sigma_{12}}{|\Sigma|}\mu_1\mu_2 \\
&= (2)x_1^2 + (4)x_2^2 - 2(1/2)x_1x_2 + x_1\Bigg( - 2(2)\mu_1 + 2(1/2)\mu_2\Bigg) + x_2\Bigg( - 2(4)\mu_2 + 2(1/2)\mu_1\Bigg) + (2)\mu_1^2 + (4)\mu_2^2 - 2(1/2)\mu_1\mu_2
\end{aligned}$$

Now we know that

$$\begin{aligned}
-11 &= -2(2)\mu_1 + 2(1/2)\mu_2 \text{ and} \\
-5 &= -2(4)\mu_2 + 2(1/2)\mu_1
\end{aligned}$$

giving us that $\mu_1 = 3$ and $\mu_2 = 1$. Note that

$$\begin{aligned}
&(2)\mu_1^2 + (4)\mu_2^2 - 2(1/2)\mu_1\mu_2 \\
&= (2)(3)^2 + (4)(1)^2 - (3)(1) \\
&= 18 + 4 - 3 \\
&= 19
\end{aligned}$$

We now know what $\Sigma$ is too:

$$\Sigma = \frac{1}{|\Sigma|} \begin{bmatrix} 4 & -1/2 \\ -1/2 & 2 \end{bmatrix}$$

The determinant of $\Sigma$ is

$$(4)(2) - (-1/2)(-1/2) = 8 - \frac{1}{4} = \frac{32}{4} - \frac{1}{4} = \frac{31}{4}$$

Now we can solve for the variance of $3X_1 - X_2$ as

$$\begin{bmatrix}3 & -1\end{bmatrix}\begin{bmatrix}\frac{16}{31} & \frac{-2}{31}\\ \frac{-2}{31} & \frac{8}{31}\end{bmatrix}\begin{bmatrix}3\\-1\end{bmatrix} = \frac{164}{31}$$

We know the distribution of the transformed variable is $N(8, \sqrt{164/31})$, and $P(3X_1 - X_2 < a) = 0.01$ is true when $a = 2.649237$.

## Problem 5.5

[Was not able to finish. I'll send it later today.]

(a) Show that $(X_1, X_2)$ has a bivariate normal distribution with means $\mu_1, \mu_2$, variances $\sigma_1^2, \sigma_2^2$, and a correlation coefficient $\rho$ if and only if every linear combination $c_1X_1 + c_2X_2$ has a univariate normal distribution with mean $c_1\mu_1 + c_2\mu_2$, and variance $c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2c_1c_2\rho\sigma_{1,2}$, where $c_1$ and $c_2$ are real constants, not both equal to zero.

---

(b) Let $Y_i = X_i/\sigma_i, i = 1, 2$. Show that $Var(Y_1 - Y_2) = 2(1 - \rho)$.

## Problem 5.6

(a) Let $(X_1, X_2) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ where $\mu_1 = \mu_2 = 0$ and $\rho \neq 1$. The polar coordinate transformation is defined by $X_1 = Rcos\Theta$, $X_2 = Rsin\Theta$. Show that the joint pdf of $R$ and $\Theta$ is given by

$$r(2\pi)^{-1}(1-\rho^2)^{-1/2}exp\left[-\frac{1}{2(1-\rho^2)}r^2(1-\rho sin2\theta)\right],$$

$0 \leq r < \infty$, and $0 \leq \theta \leq 2\pi$, and that the marginal pdf of $\Theta$ is

$$(2\pi)^{-1}(1-\rho^2)^{1/2}(1-\rho sin2\theta)^{-1},\ 0 \leq \theta \leq 2\pi.$$

---

To complete the transformation, we need to find the Jacobian:

$$\begin{aligned}
J = \begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix} &= \cos\theta r\cos\theta - (-r\sin\theta\sin\theta)\\
&= r\cos^2\theta + r\sin^2\theta\\
&= r(\cos^2\theta + \sin^2\theta)\\
&= r
\end{aligned}$$

The bivariate normal distribution is

$$\begin{aligned}
f(x_1, x_2) &= \frac{1}{2\pi|\Sigma|^{1/2}}\exp\left\{-\frac{1}{2}\frac{1}{1-\rho^2}\left(\frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2} - 2\rho\frac{x_1x_2}{\sigma_1\sigma_2}\right)\right\}\\
&= \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{1}{1-\rho^2}\left(\frac{x_1^2}{\sigma_1^2} + \frac{x_2^2}{\sigma_2^2} - 2\rho\frac{x_1x_2}{\sigma_1\sigma_2}\right)\right\}
\end{aligned}$$

Now we can find the distribution of the transformed variables:

$$
\begin{aligned}
f(r,\theta) &= \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{1}{1-\rho^2}\left(\frac{x_1^2}{\sigma_1^2}+\frac{x_2^2}{\sigma_2^2}-2\rho\frac{x_1x_2}{\sigma_1\sigma_2}\right)\right\}|r| \\
&= \frac{1}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{1}{1-\rho^2}\left(\frac{r^2\cos^2\theta}{\sigma_1^2}+\frac{r^2\sin^2\theta}{\sigma_2^2}-2\rho\frac{r\cos\theta r\sin\theta}{\sigma_1\sigma_2}\right)\right\}|r| \\
&= \frac{|r|}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{r^2}{1-\rho^2}\left(\frac{\cos^2\theta\sigma_2^2}{\sigma_1^2\sigma_2^2}+\frac{\sin^2\theta\sigma_1^2}{\sigma_1^2\sigma_2^2}-2\rho\frac{\cos\theta\sin\theta}{\sigma_1\sigma_2}\right)\right\} \\
&= \frac{|r|}{2\pi\sqrt{\sigma_1^2\sigma_2^2(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{r^2}{1-\rho^2}\left(\frac{\cos^2\theta\sigma_2^2+\sin^2\theta\sigma_1^2-\rho\sin 2\theta}{\sigma_1\sigma_2}\right)\right\}
\end{aligned}
$$

The only way for this to work is if we assume $\sigma_1=\sigma_2=1$. Then we have shown equality:

$$
\begin{aligned}
&= \frac{|r|}{2\pi\sqrt{(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{r^2}{1-\rho^2}\left(\cos^2\theta\ +\sin^2\theta-\rho\sin 2\theta\right)\right\} \\
&= \frac{r}{2\pi\sqrt{(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{r^2}{1-\rho^2}\left(1-\rho\sin 2\theta\right)\right\},\ \text{for } r\geq 0, 0\leq\theta\leq 2\pi
\end{aligned}
$$

Now we can find the mariginal density of $\Theta$:

$$
\begin{aligned}
f_\Theta(\theta) &= \int_0^\infty \frac{r}{2\pi\sqrt{(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{r^2}{1-\rho^2}\left(1-\rho\sin 2\theta\right)\right\}dr \\
&\text{Let } u=r^2:\ du=2rdr,\ rdr=\frac{du}{2} \\
&= \int_0^\infty \frac{1}{2\pi\sqrt{(1-\rho^2)}}\exp\left\{-\frac{1}{2}\frac{u}{1-\rho^2}\left(1-\rho\sin 2\theta\right)\right\}\frac{1}{2}du \\
&= \int_0^\infty \frac{1}{2\pi\sqrt{(1-\rho^2)}}\frac{\frac{1}{2}\frac{1}{(1-\rho^2)}(1-\rho\sin 2\theta)}{\frac{1}{2}\frac{1}{(1-\rho^2)}(1-\rho\sin 2\theta)}\exp\left\{-\frac{1}{2}\frac{u}{1-\rho^2}\left(1-\rho\sin 2\theta\right)\right\}\frac{1}{2}du \\
&= \frac{1}{2}\frac{1}{2\pi\sqrt{(1-\rho^2)}}\frac{1}{\frac{1}{2}\frac{1}{(1-\rho^2)}(1-\rho\sin 2\theta)}\int_0^\infty \frac{1}{2}\frac{1}{(1-\rho^2)}(1-\rho\sin 2\theta)\exp\left\{-\frac{1}{2}\frac{u}{1-\rho^2}\left(1-\rho\sin 2\theta\right)\right\}du
\end{aligned}
$$

We recognize now that we have an exponential distribution integrated over all values, thus summing to one. Now we have:

$$
\begin{aligned}
f_\Theta(\theta) &= \frac{1}{2}\frac{1}{2\pi\sqrt{(1-\rho^2)}}\frac{2(1-\rho^2)}{(1-\rho\sin 2\theta)} \\
&= \frac{1}{2\pi}\frac{\sqrt{(1-\rho^2)}}{(1-\rho\sin 2\theta)},\ 0\leq\theta\leq 2\pi
\end{aligned}
$$

(b) Suppose $(X_1,X_2)$ has a bivariate normal distribution $N_2(0,0,\sigma_1^2,\sigma_2^2,\rho,\ |\rho|\neq 1$. Show that

$$P(X_1>0,X_2>0)=\frac{1}{4}+\frac{1}{2\pi}sin^{-1}(\rho).$$

---

From part (a) we know that $0\leq\theta\leq\pi/2$, and since $X_1$ and $X_2$ are fuctions of $\theta$, $P(X_1>0,X_2>0)$ is equivalent to $P(0 \leq \theta \leq \pi/2)$. Using this information we can manipulate the density function of $\Theta$:

$$\begin{aligned}
P(0 \leq \theta \leq \pi/2) &= \int_0^{\pi/2} f(\theta) d\theta \\
&= \int_0^{\pi/2} \frac{1}{2\pi} \frac{\sqrt{1-\rho^2}}{1-\rho \sin 2\theta} d\theta \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \frac{1}{1-2\rho \sin\theta \cos\theta} d\theta \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \frac{\sec^2\theta}{\sec^2\theta} \frac{1}{1-2\rho \sin\theta \cos\theta} d\theta \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \frac{\sec^2\theta}{\sec^2\theta - 2\sec^2\theta \rho \sin\theta \cos\theta} d\theta \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \frac{\sec^2\theta}{\sec^2\theta - 2\rho \tan\theta} d\theta
\end{aligned}$$

We can let $u = \tan\theta$, $du = \sec^2\theta d\theta$. Note that our bounds have changed.

$$\begin{aligned}
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \frac{1}{1+u^2-2\rho u} du \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \frac{1}{1+u^2-2\rho u+\rho^2-\rho^2} du \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \frac{1}{(1-\rho^2)+(u-\rho)^2} du \\
&= \frac{\sqrt{1-\rho^2}}{2\pi} \int_0^{\pi/2} \frac{\frac{1}{1-\rho^2}}{1+\frac{(u-\rho)^2}{(1-\rho^2)}} du \\
&= \frac{\sqrt{1-\rho^2}}{2\pi(1-\rho^2)} \int_0^{\pi/2} \frac{\frac{1}{1-\rho^2}}{1+\left(\frac{u-\rho}{\sqrt{(1-\rho^2)}}\right)^2} du
\end{aligned}$$

Now, let $w = \frac{u-\rho}{\sqrt{1-\rho^2}}$ such that $dw = \frac{1}{\sqrt{1-\rho^2}} du$ and $\sqrt{(1-\rho^2)} dw = du$. Once again, note that the bounds of integration changes as a result of the substitution:

$$\begin{aligned}
&= \frac{1}{2\pi} \frac{1-\rho^2}{\sqrt{1-\rho^2}} \int_{-\frac{\rho}{\sqrt{1-\rho^2}}}^{\infty} \frac{1}{1+w^2} dw \\
&= \frac{1}{2\pi} \sqrt{1-\rho^2} \Big[ \tan^{-1}(w) \Big]_{-\frac{\rho}{\sqrt{1-\rho^2}}}^{\infty} \\
&= \frac{1}{2\pi} \sqrt{1-\rho^2} \Big[ \tan^{-1}(w) \Big]_{-\frac{\rho}{\sqrt{1-\rho^2}}}^{\infty} \\
&= \frac{1}{2\pi} \sqrt{1-\rho^2} \left[ \tan^{-1}(\infty) - \tan^{-1}\left( - \frac{\rho}{\sqrt{1-\rho^2}} \right) \right] \\
&= \frac{1}{2\pi} \frac{\sqrt{1-\rho^2}}{\sqrt{1-\rho^2}} \Big[ \pi/2 - \sin(-\rho) \Big] \\
&= \frac{1}{2\pi} \Big[ \pi/2 - \sin(\rho) \Big]
\end{aligned}$$

## Problem 5.7

The random vector $x = (X_1, X_2, \ldots, X_k)'$ is said to have a symmetric multivariate normal distribution if $x \sim N_k(\mu, \Sigma)$ where $\mu = \mu 1_k$, i.e., the mean of each $X_j$ is equal to the same constant $\mu$, and $\Sigma$ is the equicorrelation dispersion matrix, i.e.,

$$\Sigma = \sigma^2 = \begin{bmatrix} 1 & \rho & \ldots & \rho \\ \rho & 1 & \ldots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \ldots & 1 \end{bmatrix}$$

When $k = 3$, $\mu = 0$, $\sigma^2 = 2$, and $\rho = 1/2$, find the probability that $X_3 = min(X_1, X_2.X_3)$.

(*Hint:* Recall that if $x = (X_1, \ldots, X_k)'$ has a continuous symmetric distribution, then all possible permutations of $X_1, \ldots, X_k$ are equally likely, each having probability $P(X_{i1} < \cdots < X_{ik}) = 1/k!$ for any permutation $(i1, \ldots, i_k)$ for the first $k$ positive integers.

---

If all possible permutations of $X_1, \ldots, X_k$ are equally likely, then any event involving $X_1, X_2$, and $X_3$ are equally likely. We can write out all possible outcomes since there are only three samples:

$$\begin{gathered} X_1 < X_2 < X_3 \\ X_1 < X_3 < X_2 \\ X_2 < X_1 < X_3 \\ X_2 < X_2 < X_1 \\ X_3 < X_1 < X_2 \\ X_3 < X_2 < X_1 \end{gathered}$$

Of the six possible outcomes, $X_3$ is the minimum in two cases. Therefore, $P(X_3 = \min(X_1, X_2, X_3)) = \frac{1}{3}$.

## Problem 5.8

Let $\mathbf{x} \sim N_k(0, \Sigma)$ with pdf $f(x)$ where $\Sigma = \{\Sigma_{ij}\}$. The entropy $h(x)$ is defined as

$$h(x) = -\int f(x) ln f(x)$$

(a) Show that $h(x) = \frac{1}{2} ln\big[(2\pi e)^k |\Sigma|\big]$.

---

We need to show that

$$\frac{1}{2}\ln(2\pi)^k|\Sigma| = -\int f(x)\ln(f(x))$$

But we know $\mathbf{x}$ has a multivariate normal distribution $N_k(0, \Sigma)$:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\bigg\{-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\bigg\},\ \mathbf{x} \in R^k$$

Inserting into the definition of entropy, we have

$$\begin{aligned} &-\int \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\bigg\{-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\bigg\} \ln\bigg(\frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\bigg\{-\frac{1}{2}(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\bigg\}\bigg) \\ &= -\int \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\bigg\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x}+\mathbf{x}'\mu)\bigg\} \ln\bigg(\frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\bigg\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x}+\mathbf{x}'\mu)\bigg\}\bigg) \\ &= -\int \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\bigg\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x})\bigg\} \ln\bigg(\frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\bigg\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x})\bigg\}\bigg) \end{aligned}$$

since we were given that $\mu = \mathbf{0}$. Applying log rules, we can expand the last term and continue:

$$
\begin{aligned}
&= -\int \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x})\right\}\left(-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x}) - \frac{k}{2}\ln(2\pi) - \frac{1}{2}\ln(|\Sigma|)\right) \\
&= \int \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x})\right\}\left(\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x}) + \frac{k}{2}\ln(2\pi) + \frac{1}{2}\ln(|\Sigma|)\right) \\
&= \frac{1}{2}\int \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x})\right\}\left((\mathbf{x}'\Sigma^{-1}\mathbf{x}) + k\ln(2\pi) + \ln(|\Sigma|)\right) \\
&= \frac{1}{2}\left[\frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}}\int (\mathbf{x}'\Sigma^{-1}\mathbf{x}) \exp\left\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x})\right\}\right. \\
&\quad + k\ln(2\pi)\int \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x})\right\} \\
&\quad \left. + \ln(|\Sigma|)\int \frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}'\Sigma^{-1}\mathbf{x})\right\}\Big)\right]
\end{aligned}
$$

The integrals in the last two terms simply equal the constants pulled out front as we are integrating the multinomial normal distribution across all values of $\mathbf{x}$. The first term can be evaluated per Result 5.1.3:

$$
\begin{aligned}
&\frac{1}{2}\left[\frac{1}{(2\pi)^{k/2}|\Sigma|^{1/2}}(2\pi)^{k/2}|\Sigma^{-1}|^{-1/2}tr(\Sigma^{-1}\Sigma) + k\ln(2\pi) + \ln(|\Sigma|)\right] \\
&= \frac{1}{2}\left[\frac{1}{|\Sigma|^{1/2}}|\Sigma|^{1/2}tr(I_k) + k\ln(2\pi) + \ln(|\Sigma|)\right] \\
&= \frac{1}{2}\left[k + k\ln(2\pi) + \ln(|\Sigma|)\right] \\
&= \frac{1}{2}\left[k + \ln\left[(2\pi)^k|\Sigma|\right]\right] \\
&= \frac{1}{2}\left[\ln(e^k) + \ln\left[(2\pi)^k|\Sigma|\right]\right] \\
&= \frac{1}{2}\ln\left[(2\pi e)^k|\Sigma|\right]
\end{aligned}
$$

(b) Hence, or otherwise, show that $|\Sigma| \leq \prod_{i=1}^{k}\Sigma_{ii}$, with equality holding if and only if $\Sigma_{ij} = 0$, for $i \neq j$ [Hadamard's inequality].

From part (a) we know that $-\int f(x)\ln(f(x)) = \frac{1}{2}\ln\left[(2\pi e)^k|\Sigma|\right]$.

Let's consider what happens when we let $k = 2$ and see if we can generalize from there. From tohe prompt, we know that

$$
h(x_1, x_2) = -\int f(x)\ln(f(x)) = -E[\ln(f(X))] = E[-\ln(f(X))]
$$

Consider what happens when we add information about $X_1$ to the entropy of $X_2$:

$$
\begin{aligned}
h(X_1) + h(X_2|X_1) &= E[-\ln(f(x_1))] + E[-\ln(f(x_2|x_1))] \\
&= E[-\ln(f(x_1))f(x_2|x_1)] \\
&= E[-\ln(f(X))] \\
&= h(x_1, x_2)
\end{aligned}
$$

That is, total uncertainty can be partitioned into $k$ entropies for all $k$.

What happens when we subtract conditional entropy?

$$\begin{aligned}
h(x_2) - h(x_2|x_1) &= E[-\ln(f(x_2))] - E[-\ln(f(x_2|x_2))] \\
&= -E[\ln(\frac{f(x_2)}{f(x_2|x_2)})] \\
&= -E[\ln(\frac{f(x_2)}{\frac{f(x_1,x_2)}{f(x_1)}})] \\
&= -E[\ln(\frac{f(x_2)f(x_1)}{f(x_1,x_2)})]
\end{aligned}$$

We know from Jensen's Inequality that

$$\begin{aligned}
E\left[\ln\left(\frac{f(x_1)f(x_2)}{f(x_1,x_2)}\right)\right] &\leq \ln\left(E\left[\frac{f(x_1)f(x_2)}{f(x_1,x_2)}\right]\right) \\
&= \ln\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f(x_1)f(x_2)}{f(x_1,x_2)}f(x_1,x_2)dx_1dx_2\right) \\
&= \ln\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x_1)f(x_2)dx_1dx_2\right) \\
&= \ln(1) = 0
\end{aligned}$$

This implies that

$$E\left[\ln\left(\frac{f(x_1)f(x_2)}{f(x_1,x_2)}\right)\right] \leq 0$$

and also that

$$-E\left[\ln\left(\frac{f(x_1)f(x_2)}{f(x_1,x_2)}\right)\right] \geq 0$$

Recall, however, that

$$\begin{aligned}
&h(x_2) - h(x_2|x_1) = -E\left[\ln\left(\frac{f(x_1)f(x_2)}{f(x_1,x_2)}\right)\right] \\
&\implies h(x_2) - h(x_2|x_1) \geq 0 \\
&\implies h(x_2) \geq h(x_2|x_1)
\end{aligned}$$

Let's combine this result with what found when we added the conditional entropy:

$$\begin{aligned}
&h(x_1,x_2) = h(x_1) + h(x_2|x_1) \\
&\implies h(x_1,x_2) - h(x_1) = h(x_2|x_1) \\
&\implies h(x_1,x_2) - h(x_1) \leq h(x_2) \\
&\implies h(x_1,x_2) \leq h(x_1) + h(x_2)
\end{aligned}$$

Generalizing this, we see that

$$h(x_1,\ldots,h_k) \leq \sum_{i=1}^{k} h(x_i)$$

From part (a) we know that

$$h(x_1,\ldots,h_k) = \frac{1}{2}\ln[(2\pi e)^k|\Sigma|]$$

We can also expand the summation of component entropies:

$$\begin{aligned}\sum_{i=1}^{k} h(x_i) &= h(x_1) + h(x_2) + \cdots + h(x_k) \\ &= \frac{1}{2}\ln[(2\pi e)^k |\Sigma_{11}|] + \frac{1}{2}\ln[(2\pi e)^k |\Sigma_{22}|] + \cdots + \frac{1}{2}\ln[(2\pi e)^k |\Sigma_{kk}|] \\ &= \frac{1}{2}\ln\left[(2\pi e)^k \prod_{i=1}^{k} |\Sigma_{ii}|\right]\end{aligned}$$

Now we can make comparisons to total entropy:

$$\begin{aligned}&\sum_{i=1}^{k} h(x_i) = \frac{1}{2}\ln\left[(2\pi e)^k \prod_{i=1}^{k} |\Sigma_{ii}|\right] \\ &\implies h(x_1, \ldots, x_k) \leq \frac{1}{2}\ln\left[(2\pi e)^k \prod_{i=1}^{k} |\Sigma_{ii}|\right] \\ &\implies \frac{1}{2}\ln[(2\pi e)^k |\Sigma|] \leq \frac{1}{2}\ln\left[(2\pi e)^k \prod_{i=1}^{k} |\Sigma_{ii}|\right] \\ &\implies \ln[(2\pi e)^k |\Sigma|] \leq \ln\left[(2\pi e)^k \prod_{i=1}^{k} |\Sigma_{ii}|\right] \\ &\implies (2\pi e)^k |\Sigma| \leq (2\pi e)^k \prod_{i=1}^{k} |\Sigma_{ii}| \\ &\implies |\Sigma| \leq \prod_{i=1}^{k} |\Sigma_{ii}|\end{aligned}$$