Homework 6 Solutions

1. ***A Hovercraft moving under random wind gusts***

Here we will consider a hovercraft with mass 1 moving under a random wind gusts. We will use an extremely simplified model of wind gusts. We'll see how to improve this model later in the course.

We have a hovercraft moving in the plane with two thrusters, each pointing through the center of mass, exerting forces in the **x** and **y** directions. The hovercraft has mass 1. The discretized equations of motion for the hovercraft are

$$x(t+1) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} \frac{1}{2} & 0 \\ 1 & 0 \\ 0 & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

where x_1 and x_2 are the position and velocity in the **x**-direction, and x_3, x_4 are the position and velocity in the **y**-direction. Here

$$u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

is the force acting on the hovercraft for time in the interval $[t, t+1)$.

Suppose the forces acting on the hovercraft are i.i.d. and Gaussian, with $u(t) \in \mathcal{N}(0, \Sigma_u)$, and $u(s)$ and $u(t)$ uncorrelated when $s \neq t$, and

$$\Sigma_u = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

(a) Find an expression for the 90% confidence ellipsoid for the position of the hovercraft after 10 seconds. Plot the ellipsoid in $\mathbb{R}^2$.

(b) Simulate the system 1000 times and plot the resulting final positions in the plane. Count (using Matlab) the percentage of final positions that land in the 90% confidence ellipsoid.

(c) Also plot the 90%, 70% and 50% confidence ellipsoids on the same plot.

(d) The final position $x \in \mathbb{R}^2$ is Gaussian, with mean zero. Suppose its covariance is Σ. What is the pdf of $x^T \Sigma^{-1} x$? Compute $x^T \Sigma^{-1} x$ for 1000 simulations and plot the empirical cdf on top of the actual cdf.

Solution.

(a) The position of the hovercraft at time t is $y(t)$, where

$$y(t) = Cx(t)$$

and

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

We can express $y(t)$ directly in terms of the inputs as

$$\begin{aligned} y(t) &= CA^t x(0) + \sum_{k=0}^{t-1} CA^{t-k-1} B u(k) \\ &= CA^t x(0) + H u_{\text{seq}} \end{aligned}$$

where

$$u_{\text{seq}} = \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(t-1) \end{bmatrix} \qquad H = \begin{bmatrix} CA^{t-1}B & CA^{t-2}B & \dots & CAB & CB \end{bmatrix}$$

Since the inputs are i.i.d. with $u(t) \sim \mathcal{N}(0, \Sigma_u)$ for all t, we have

$$u_{\text{seq}} \sim \mathcal{N}(0, \Sigma)$$

where

$$\Sigma = \begin{bmatrix} \Sigma_u & & & \\ & \Sigma_u & & \\ & & \ddots & \\ & & & \Sigma_u \end{bmatrix}$$

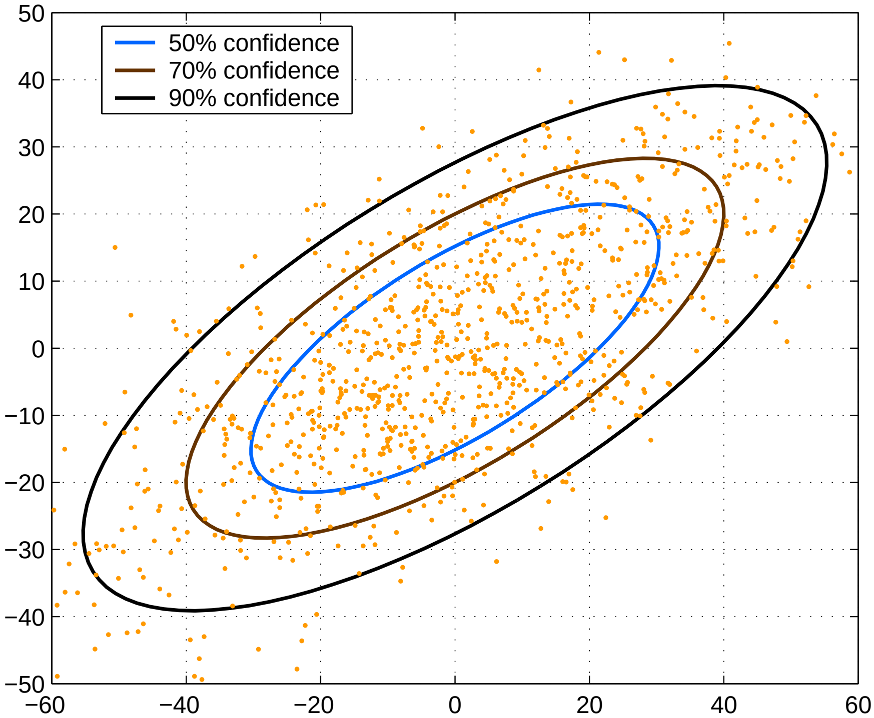
Now $y(t)$ is an affine function of u_{seq}, so it is Gaussian, and we have $y(t) \sim \mathcal{N}(\mu_y, \Sigma_y)$ where

$$\mu_y = CA^t x(0) \qquad \Sigma_y = H \Sigma H^T$$

The 90% confidence ellipsoid for $y(t)$ is

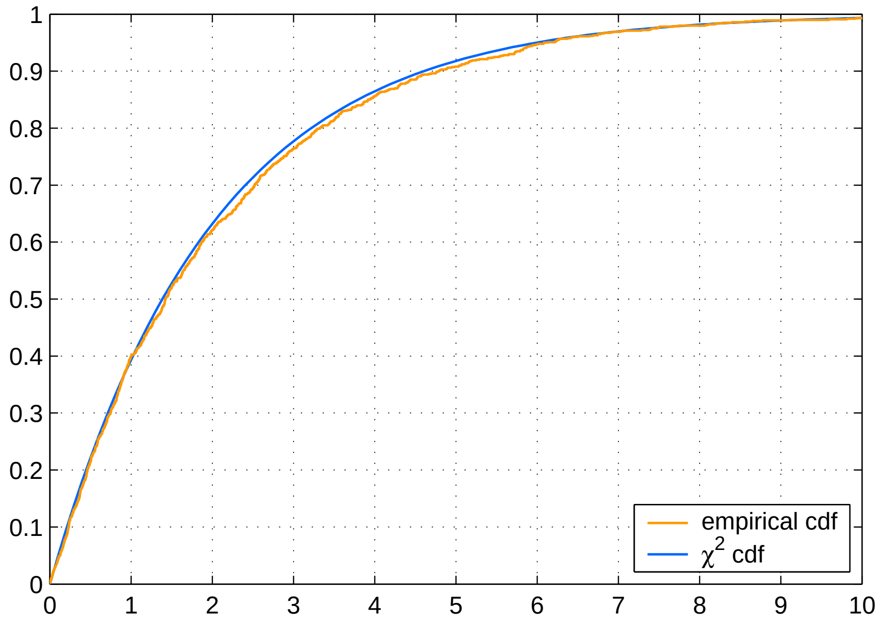
$$\left\{ z \in \mathbb{R}^2 \mid \left(z - CA^t x(0)\right)^T \Sigma_y^{-1} \left(z - CA^t x(0)\right) \leq F_{\chi_2^2}^{-1}(0.9) \right\}$$

(b) The sampled points are shown below.



(c) The confidence ellipsoids are shown above.

(d) The random variable $q = y(t)^T \Sigma_y^{-1} y(t)$ has a χ_2^2 distribution. The empirical cdf and the actual cdf are plotted below.



2. ***The Chebyshev inequality for vectors and confidence ellipsoids***

(a) Suppose $x : \Omega \to \mathbb{R}$. Show that the Chebyshev inequality can be written as

$$\text{Prob}\left(\frac{(x-\mu)^2}{\text{cov}(x)} \geq a\right) \leq \frac{1}{a}$$

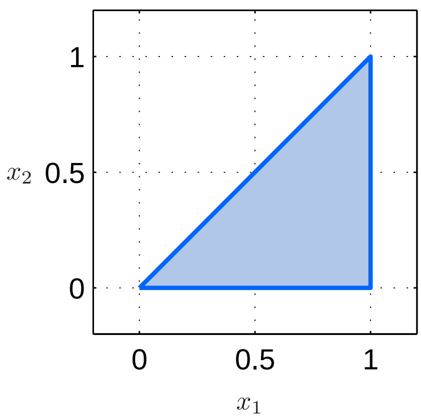
where $\mu = \mathrm{E}\, x$.

(b) Suppose $x : \Omega \to R$ is a scalar random variable, which is uniformly distributed on the interval $[-1, 1]$. The Chebyshev bound gives a lower bound that x lies in the interval $[-c, c]$. What are the values of c corresponding to lower bounds of 0.1, 0.5 and 0.9?

(c) Now let's consider a random vector $x : \Omega \to \mathbb{R}^n$, with mean $\mathrm{E}\, x = \mu$ and covariance matrix $\text{cov}(x) = \Sigma$. Generalize the Chebyshev inequality to find an upper bound on the probability

$$\text{Prob}\Big((x-\mu)^T \Sigma^{-1}(x-\mu) \geq a\Big)$$

(d) This can be used to construct *confidence ellipsoids* for pdfs which are not Gaussian. For example, consider the induced pdf for the random variable $x : \Omega \to \mathbb{R}^2$ which is uniform on the triangle shown below.



What is the mean and covariance of x?

(e) Using the Chebyshev bound you derived above, plot ellipsoids E_α such that

$$\mathrm{Prob}(x \in E_\alpha) \geq \alpha$$

for the cases where α is 0.1, 0.5 and 0.9.

Solution.

(a) We have the Chebyshev inequality

$$\mathrm{Prob}(|x-\mu| \geq a) \leq \frac{\mathrm{cov}(x)}{a^2}$$

which is equivalent to

$$\mathrm{Prob}((x-\mu)^2 \geq a^2) \leq \frac{\mathrm{cov}(x)}{a^2}$$

which is also equivalent to

$$\mathrm{Prob}\left(\frac{(x-\mu)^2}{\mathrm{cov}(x)} \geq \frac{a^2}{\mathrm{cov}(x)}\right) \leq \frac{\mathrm{cov}(x)}{a^2}$$

Now letting $b = a^2/\operatorname{cov}(x)$ gives the desired result.

(b) Since $x \sim \mathcal{U}[-1,1]$, integrating gives

$$\mathrm{cov}(x) = \frac{1}{3}$$

Also the Chebyshev bound is

$$\mathrm{Prob}(|x| \leq c) \geq p$$

where

$$p = 1 - \frac{\mathrm{cov}(x)}{c^2}$$

Hence

$$c = \sqrt{\frac{\mathrm{cov}(x)}{1-p}}$$

For $p = 0.1$, $c = \sqrt{\frac{10}{27}} \approx 0.609$, for $p = 0.5$, $c = \sqrt{\frac{2}{3}} \approx 0.816$ and for $p = 0.9$ we have $c = \sqrt{\frac{10}{3}} \approx 1.83$. Note of course that the latter bound is trivial, since all probabilities are less than or equal to one.

(c) For any function $f : \mathbb{R}^n \to \mathbb{R}$ such that $f(x) \geq 0$ for all x the Markov inequality implies that

$$\mathrm{Prob}\Big(x \in \Big\{\, y \in \mathbb{R}^n \mid f(y) \geq a \,\Big\}\Big) \leq \frac{1}{a}\,\mathrm{E}\, f(x)$$

We'll use

$$f(x) = (x-\mu)^T \Sigma^{-1}(x-\mu)$$

We have

$$\begin{aligned}
\mathrm{E}\, f(x) &= \mathrm{E}(x-\mu)^T\Sigma^{-1}(x-\mu) \\
&= \mathrm{E}\,\mathrm{trace}\Big((x-\mu)^T\Sigma^{-1}(x-\mu)\Big) && \text{since } f \text{ is scalar} \\
&= \mathrm{E}\,\mathrm{trace}\Big(\Sigma^{-1}(x-\mu)(x-\mu)^T\Big) && \text{since } \mathrm{trace}(AB) = \mathrm{trace}(BA) \\
&= \mathrm{trace}\Big(\Sigma^{-1}\,\mathrm{E}\big((x-\mu)(x-\mu)^T\big)\Big) \\
&= \mathrm{trace}\Big(\Sigma^{-1}\Sigma\Big) \\
&= n
\end{aligned}$$

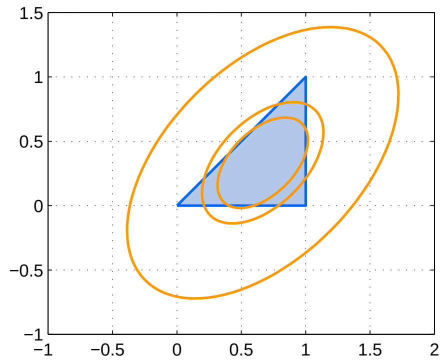
Hence

$$\mathrm{Prob}\Big((x-\mu)^T\Sigma^{-1}(x-\mu) \geq a\Big) \leq \frac{n}{a}$$

(d) By integration we find

$$\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \frac{1}{18} & \frac{1}{36} \\ \frac{1}{36} & \frac{1}{18} \end{bmatrix}$$

(e) The ellipsoids are below.



3. ***Conditional pdfs for Gaussians***

Suppose $x \in \mathbb{R}^n$ and $x \sim \mathcal{N}(0, \Sigma)$. Partition x as $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where $x_1 \in \mathbb{R}^r$, and partition Σ as

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix}$$

so that $\Sigma_{11} \in \mathbb{R}^{r \times r}$. First let's consider the special case when $n = 2$, $r = 1$ and

$$\Sigma = \begin{bmatrix} 2 & 0.8 \\ 0.8 & 1 \end{bmatrix}$$

(a) Plot the 90% confidence ellipsoid C for x.

(b) Suppose now we are given $x_2 = 1$. Find the 90% confidence interval $[c_1, c_2]$ for the conditional $x_1 | (x_2 = 1)$, and find its mean. Plot three vertical lines through the boundary points of this interval and the mean.

(c) Now collect 10000 samples of x. In Matlab, pick out those samples for which $0.9 \leq x_2 \leq 1.1$, and discard the rest. Count the percentage of those you kept for which x_1 lies in $[c_1, c_2]$.

(d) The ellipsoid C intersects the line $x_2 = 1$ at points x_1 in the interval $[c_3, c_4]$. Find this interval.

(e) More generally, suppose E is the ellipsoid

$$E = \left\{ x \in \mathbb{R}^n \mid x^T \Sigma^{-1} x \leq 1 \right\}$$

The intersection of E with the surface

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^n \mid x_2 = b \right\}$$

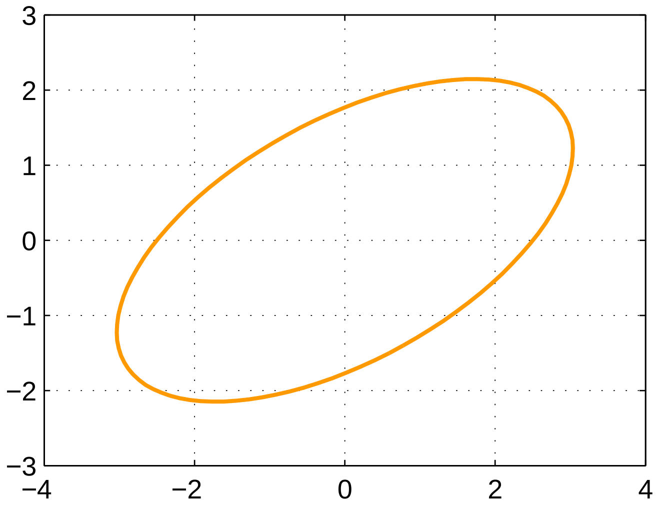
is an ellipsoid, which has the form

$$F = \left\{ x_1 \in \mathbb{R}^r \mid (x_1 - a)^T Q^{-1} (x_1 - a) \leq c \right\}$$

Find a, c and Q in terms of b and Σ.

Solution.

(a) The confidence ellipsoid is below.



(b) The conditional pdf of $x_1 \mid x_2 = b$ is Gaussian, with

$$a = \mathrm{E}(x_1 \mid x_2 = b) = \Sigma_{12}\Sigma_{22}^{-1}b$$
$$Q = \mathrm{cov}(x_1 \mid x_2 = b) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

The conditional confidence interval for x is $[a - q, a + q]$, where

$$q = \sqrt{\beta Q}$$

and $F_{\chi_1^2}(\beta) = 0.9$. In this case $q \approx 1.918$ and $a = 0.8$. The confidence ellipsoid with the bars is below.

(c) The percentage is about 90%.

(d) From the next part, the interval is $[a - h, a + h]$ where the half-width h is

$$h = \left(Q(\alpha - b^T\Sigma_{22}^{-1}b)\right)^{\frac{1}{2}}$$

where $F_{\chi_2^2}(\alpha) = 0.9$ and $b = 1$. In this case, $h \approx 2.214$.

(e) The intersection of E with $\{ x \in \mathbb{R}^n \mid x_2 = a \}$ is the set of x_1 such that

$$\begin{bmatrix} x_1 \\ b \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} x_1 \\ b \end{bmatrix} \leq 1$$

Completing the square gives

$$\Sigma^{-1} = \begin{bmatrix} I & 0 \\ \Sigma_{22}^{-1}\Sigma_{21} & \end{bmatrix} \begin{bmatrix} Q^{-1} & 0 \\ 0 & \Sigma_{22}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{bmatrix}$$

and hence the intersection is the set of x_1 such that

$$(x_1 - a)^T Q^{-1}(x_1 - a) \leq 1 - b^T \Sigma_{22}^{-1} b$$

where Q and a are as in part (b).