16.09 Probability and Statistics
Spring 2020 Project

Prof. Youssef Marzouk (ymarz@mit.edu)
Karthik Gopalakrishnan (karthikg@mit.edu)
Massachusetts Institute of Technology

29 April 2020

Preliminaries

- **Due date:** 10 pm EDT on Tuesday 12 May 2020
- Team size: Up to 3 students
- This is an open-ended project. Please make any assumptions, design choices, or algorithmic choices as needed and justify them. There may be more than one correct way to approach a problem.
- Deliverables: Please upload to Stellar *one joint report* for the group, in PDF format, along with all the codes that were used. Please don't spend too much time cleaning up the code; we just want to get a sense of your implementation and assign partial credit if applicable. At the beginning of the report, describe the contribution of each team member towards the project.
- Grading rubric: Each problem is worth 25 points.
- If you would like to discuss any aspect of the project, please use Piazza, email us, or attend our office hours.
- Most importantly, remember that the aim of the project is to help you learn the subject content, and to provide a more relaxed way of internalizing probability and statistics while working in smaller groups. So if anything is challenging, talk to someone (or the instructors) and don't get stressed.
- You may use any result from the class, the course textbook, or any online/offline resource directly as long as you understand the material and provide appropriate references. Remember: you are responsible for all the content that is presented in the final report.

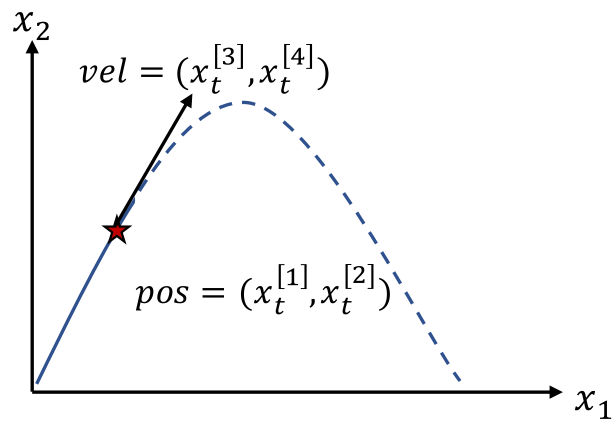

Figure 1: Setup showing the position and velocity of the projectile at time t

The projectile tracking problem

The goal of this project is to develop techniques to track the motion of a projectile based on radar measurements, and to quantify the uncertainty in this motion. We will apply concepts from probability, statistics, and control systems to approach this problem.

Let us consider the motion of a projectile in a two-dimensional plane, and construct a discrete-time state space representation of its motion.

The state of the system at time t is captured by specifying the position and velocity of the projectile. Thus, $\boldsymbol{x}_t \coloneqq (x_t^{[1]}, x_t^{[2]}, \mathrm{d}x_t^{[1]}/\mathrm{d}t, \mathrm{d}x_t^{[2]}/\mathrm{d}t)$ is the state of the system at t, consisting of the projectile coordinates $(x_t^{[1]}, x_t^{[2]})$ and their time derivatives. The subscript t is always a time index. We will think of the coordinate $x^{[1]}$ as a horizontal position and the coordinate $x^{[2]}$ as a vertical position, i.e., aligned with gravity. For simplicity, we will write the velocities as additional numbered components of the state, that is, $x_t^{[3]} \coloneqq \mathrm{d}x_t^{[1]}/\mathrm{d}t$ and $x_t^{[4]} \coloneqq \mathrm{d}x_t^{[2]}/\mathrm{d}t$.

In flight, the projectile is subject to aerodynamic forces (lift and drag) and the force of gravity. In this project, we will ignore any systematic contributions to lift and drag, and consider only the effect of gravity and random atmospheric disturbances (e.g., wind gusts)

In such a case, the discrete-time evolution of the state of the projectile can be written succinctly as

$$\boldsymbol{x}_{t+1} = \begin{bmatrix} x_{t+1}^{[1]} \\ x_{t+1}^{[2]} \\ x_{t+1}^{[3]} \\ x_{t+1}^{[4]} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_t^{[1]} \\ x_t^{[2]} \\ x_t^{[3]} \\ x_t^{[4]} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -g\Delta t \end{bmatrix} + \boldsymbol{w}_t \tag{1}$$

where Δt is the discrete time interval, $g = 9.8$ m/s^2 is the acceleration due to gravity, and $\boldsymbol{w}_t \in \mathbb{R}^4$ is a multivariate Gaussian noise term with mean $\mathbf{0} \in \mathbb{R}^4$ and covariance $Q \in \mathbb{R}^{4\times 4}$. The time is discretized as $\{1, 2, \ldots, t, t+1, \ldots\}$, where the time difference between any two successive indices is Δt.

We monitor the trajectory of the projectile using our ground-based radar, which outputs the position of the projectile. The radar is not a perfect instrument, and its measurement error at every time t (in both the horizontal and vertical directions) is modelled as a bivariate Gaussian $\boldsymbol{r}_t \sim \mathcal{N}(0, R)$. Thus the output from a radar at time t, $\boldsymbol{y}_t \in \mathbb{R}^2$ is given by

$$\boldsymbol{y}_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \boldsymbol{x}_t + \boldsymbol{r}_t \tag{2}$$

where the measurement errors are independent at different times, i.e., $\boldsymbol{r}_t$ and $\boldsymbol{r}_s$ are independent for $t \neq s$.

Thus, given an initial position and velocity of the projectile, i.e., $\boldsymbol{x}_0$, we can model the evolution of the state using Equation (1). The observation from the radar at any time instant t, $\boldsymbol{y}_t$, is dependent on the underlying state of the system $\boldsymbol{x}_t$ and is obtained through Equation (2).

To summarize, we emphasize that there are uncertainties in the actual motion of the projectile (due to wind and atmospheric disturbances) *and* noise in the measurement of the position using our radar. Thus, simply knowing the initial launch condition is not sufficient to deterministically predict the trajectory. Nor are any measurements from the radar a perfect estimate of the position of the target.

Your mission

The overall setup is as follows: We collect (noisy) radar measurements for a part of the projectile's trajectory. Using these measurements, we would like to predict the future landing time and landing location of the projectile. Specifically, we would like to use our knowledge of probability, statistics, and Kalman filters to make quantitative statements about the uncertainty of these predictions. The problems in the project are designed to help you sequentially build a solution towards this goal. This target tracking problem setup is very common in aerospace engineering applications—for vehicles at all scales and both in air and in space. Through this project, we hope to introduce a realistic engineering application of the concepts you've learned in 16.09!

A summary of Kalman filters

Consider a generic setup where the state of the system is given by $\boldsymbol{x}_t \in \mathbb{R}^n$, the input to the system at time t is $\boldsymbol{u}_t$, and the measurements from a sensor are $\boldsymbol{y}_t \in \mathbb{R}^m$. The evolution of the system and the measurement dynamics are given by

$$\boldsymbol{x}_{t+1} = A\boldsymbol{x}_t + B\boldsymbol{u}_t + \boldsymbol{w}_t \tag{3}$$

$$\boldsymbol{y}_t = C\boldsymbol{x}_t + \boldsymbol{r}_t \tag{4}$$

where $\boldsymbol{w}_t \sim \mathcal{N}(0, Q)$ and $\boldsymbol{r}_t \sim \mathcal{N}(0, R)$. You should be able to map our projectile dynamics and radar measurement to this setup by appropriately identifying the matrices A, B, and C and the vector $\boldsymbol{u}_t$.

In the remainder of this project description, we will use the following Bayesian perspective and notation. The posterior mean of the state $\boldsymbol{x}$ at time t_1, conditioned on observations up to time t_2 (i.e., on $\{\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t_2}\}$), is denoted by $\widehat{\boldsymbol{x}}_{t_1|t_2}$. Similarly, the posterior covariance matrix of the state at time t_1 given observations through time t_2 is denoted by $P_{t_1|t_2}$. For the problems we consider here, $t_1 \geq t_2$.

In the Kalman filter setup, we begin with a distribution on the initial condition $\boldsymbol{x}_0$. This distribution is a multivariate Gaussian with mean $\widehat{\boldsymbol{x}}_{0|0}$ and covariance $P_{0|0}$, i.e., $\boldsymbol{x}_0 \sim \mathcal{N}(\widehat{\boldsymbol{x}}_{0|0}, P_{0|0})$. Now our objective is to infer the state $\boldsymbol{x}_t$ at future times, as more and more measurements $\boldsymbol{y}_t$ are obtained. Filtering focuses on estimating the state (and its uncertainty) at the current time t given data up to and including that time. In other words, we will characterize $\boldsymbol{x}_t$ by its posterior mean $\widehat{\boldsymbol{x}}_{t|t}$ and its covariance $P_{t|t}$. Then, once the next observation $\boldsymbol{y}_{t+1}$ arrives, we will construct $\widehat{\boldsymbol{x}}_{t+1|t+1}$ and the covariance $P_{t+1|t+1}$. We want a *recursive* algorithm for performing this update from t to $t+1$. With such an algorithm in hand, we could keep learning the state for an arbitrarily long time horizon.

The two main steps in recursively estimating the state are described below. In this framework, the outputs of the filter at time $t+1$ are: (i) the posterior mean vector $\widehat{\boldsymbol{x}}_{t+1|t+1}$, which represents an estimate of the state at time $t+1$ using all measurements of $\boldsymbol{y}$ obtained up to and including time $t+1$; and (ii) the posterior covariance matrix $P_{t+1|t+1}$, which represents *uncertainty* in the value of the state vector at time $t+1$. The idea of these recursive formulas is to use the previous (time t) state estimate $\widehat{\boldsymbol{x}}_{t|t}$ and covariance $P_{t|t}$, the new measurement $\boldsymbol{y}_{t+1}$, knowledge of the system dynamics (the A, B, and C matrices), control input $\boldsymbol{u}_t$, and noise characteristics (Q and R) to obtain $\widehat{\boldsymbol{x}}_{t+1|t+1}$ and $P_{t+1|t+1}$.

The two steps of Kalman filtering are:

1. Prediction step:

$$\text{Prior mean:} \qquad \widehat{\boldsymbol{x}}_{t+1|t} = A\widehat{\boldsymbol{x}}_{t|t} + B\boldsymbol{u}_t \tag{5}$$

$$\text{Prior covariance:} \qquad P_{t+1|t} = AP_{t|t}A^\top + Q \tag{6}$$

2. Update step:

$$\text{State residual:} \qquad \tilde{\boldsymbol{y}}_{t+1} = \boldsymbol{y}_{t+1} - C\widehat{\boldsymbol{x}}_{t+1|t} \tag{7}$$

$$\text{Covariance residual:} \qquad S_{t+1} = C\,P_{t+1|t}\,C^\top + R \tag{8}$$

$$\text{Kalman gain:} \qquad K_{t+1} = P_{t+1|t}\,C^\top\,S_{t+1}^{-1} \tag{9}$$

$$\text{Updated state:} \qquad \widehat{\boldsymbol{x}}_{t+1|t+1} = \widehat{\boldsymbol{x}}_{t+1|t} + K_{t+1}\tilde{\boldsymbol{y}}_{t+1} \tag{10}$$

$$\text{Updated covariance:} \qquad P_{t+1|t+1} = (\mathbb{I} - K_{t+1}C)P_{t+1|t} \tag{11}$$

Note that $\mathbb{I}$ is an $n \times n$ identity matrix.

Problem assumptions

Now we give some assumptions specific to our projectile/radar problem. Let $\Delta t = 1$ sec and let the covariance matrices Q and R be given by

$$Q = \begin{bmatrix} 20 & 0 & 5 & 0 \\ 0 & 20 & 0 & 5 \\ 5 & 0 & 10 & 0 \\ 0 & 5 & 0 & 10 \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 9 & 3 \\ 3 & 9 \end{bmatrix}. \tag{12}$$

Assume the following mean and covariance for the initial condition:

$$\widehat{\boldsymbol{x}}_{0|0} = \begin{bmatrix} 0 \\ 0 \\ 100 \\ 100 \end{bmatrix} \quad \text{and} \quad P_{0|0} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 25 & 5 \\ 0 & 0 & 5 & 25 \end{bmatrix}. \tag{13}$$

Before you get started

(a) Reading assignment: Bivariate Gaussians (lectures on March 2 and March 9, and Recitation 5); multivariate Gaussians (Recitation 10), and Kalman filters (lecture on May 4).

(b) A MATLAB script to plot the confidence regions for a bivariate Gaussian is provided.

Problem 1: Making predictions

In this problem, we will use the information available at time $t = 0$ to make predictions about the future state of the projectile. In other words, we will not yet use any data $\boldsymbol{y}_t$.

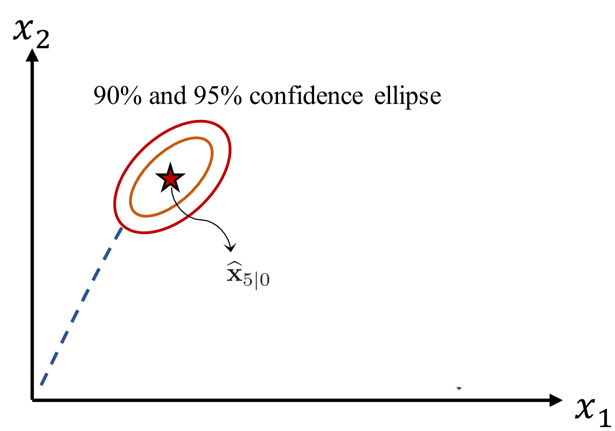

Figure 2: Example of a confidence region for problem 1

(a) What does the structure (i.e., location of the zero and non-zero elements) of the covariance matrix Q as well as the sign of the entries physically represent? Does it make sense?

(b) How would you physically interpret the covariance matrix $P_{0|0}$?

(c) Using only the prediction equations (5) and (6) of the Kalman filter (since there is no data to condition on yet), compute the mean and covariance of the projectile's state at $t = 5$. In other words, compute $\widehat{\boldsymbol{x}}_{5|0}$ and $P_{5|0}$.

(d) What are the 90% and 95% probability-enclosing ellipses for the location of the projectile at $t = 5$? Plot them as in Figure 2.

(e) Plot the 90% probability-enclosing ellipse for the position of the projectile 10 and 15 seconds after it has been thrown. Comment on the size of the ellipse, and the factors that might determine the observed trend.

Problem 2: When and where will the projectile land?

We continue from where we left off in Problem 1. Now we can (probabilistically) predict the location of the projectile at any time $t > 0$. Our probabilistic description for the projectile's location $(x_t^{[1]}, x_t^{[2]})$ at time t is a bivariate Gaussian, i.e., $\begin{pmatrix} x_t^{[1]} \\ x_t^{[2]} \end{pmatrix} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$. Here $\boldsymbol{\mu} \in \mathbb{R}^2$ and $\Sigma \in \mathbb{R}^{2\times 2}$ are the mean and covariance of a marginal of the four-dimensional Gaussian distribution of the state at time t, $\boldsymbol{x}_t \sim \mathcal{N}(\widehat{\boldsymbol{x}}_{t|0}, P_{t|0})$.[1]

In this problem, we will estimate the pmf of the landing time (since the landing time is a discrete random variable) and the pdf of the landing location (the landing location is a continuous random variable) of the projectile. Refer to Figure 3 for a more visual description of the problem.

The landing time of the projectile is solely determined by the sign of the variable $x^{[2]}$. The earliest time at which $x^{[2]} \leq 0$ is defined as the landing time for the projectile. The discrete random variable representing the landing time is denoted as T_{land}.

Let X_{land} be the random variable representing the location ($x^{[1]}$ coordinate) where the projectile lands. This location is slightly more complex to deduce. This is because of the time discretization, where at one instant the projectile may have $x^{[2]} > 0$ and at the next instant it may have a value $x^{[2]} < 0$. Thus it becomes unclear at which exact location the projectile crossed the boundary $x^{[2]} = 0$.

Because the projectile dynamics are described in discrete time, we define the landing time of a projectile as the *earliest time* at which $x^{[2]}$ becomes less than or equal to zero. We also make the following simplifying assumption: For a given (discrete) time t, the landing location is given by a *conditional* of the bivariate Gaussian above—in particular, by conditioning $x_t^{[1]}$ on the value $x_t^{[2]} = 0$. In other words, the pdf of landing location given any landing time t is taken to be:

$$f(X_{\text{land}} | T_{\text{land}} = t) \;=\; f(x_t^{[1]} | x_t^{[2]} = 0). \tag{14}$$

[1] Refer to Recitation 10 to see how we can obtain marginals from multivariate Gaussians, i.e., how to obtain $(\boldsymbol{\mu}, \Sigma)$ from $(\widehat{\boldsymbol{x}}_{t|0}, P_{t|0})$.

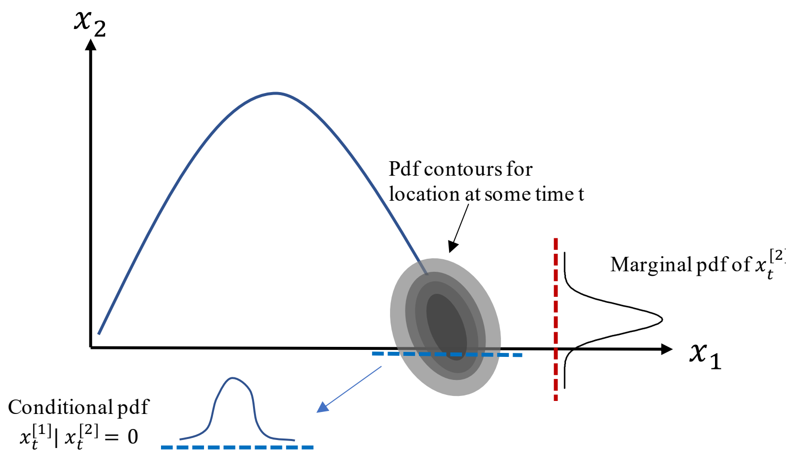

Figure 3: Schematic for Problem 2

Answer the next set of questions to derive the distributions of T_{land} and X_{land}.

(a) Compute the probability that the projectile has landed by time $t = 20$, i.e., $\mathbb{P}[T_{\text{land}} \leq 20]$. (*Hint: Compute an appropriate area under the marginal pdf of* $x^{[2]}_{t=20}$.)

(b) Using the strategy in part (a) for several different values of t to compute the CDF for T_{land}. Use this CDF to compute the pmf of T_{land}.

(c) Next, we focus on the landing location. What is the distribution of $X_{\text{land}} |\ T_{\text{land}} = 20$? Use the observation that it is identical to the distribution of $x^{[1]}_{t=20} \mid x^{[2]}_{t=20} = 0$, and is hence Gaussian. What is its mean and variance?[2]

(d) Use the strategy in part (c) to characterize the conditional distribution of $X_{\text{land}} | T_{\text{land}} = t$ for several values of t. Then, use the total probability theorem as

$$f_{X_{\text{land}}}(x) = \sum_t f_{X_{\text{land}}|T_{\text{land}}}(x \mid T_{\text{land}} = t)\, p_{T_{\text{land}}}(t) \tag{15}$$

to obtain the distribution of X_{land}. Interestingly, the distribution of X_{land} is not a Gaussian. Instead, the pdf of X_{land} is a linear combination of the densities of several different Gaussians. This distribution for X_{land} has a special name: it is called a *mixture of Gaussians*. How would you compute the mean and variance of X_{land}?

[2] Recall from lectures that if X_1 and X_2 are bivariate Gaussian with mean vector (μ_1, μ_2) and covariance matrix $[\sigma_1^2 \;\; \rho\sigma_1\sigma_2; \rho\sigma_1\sigma_2 \;\; \sigma_2^2]$, then a conditional distribution is

$$X_1|(X_2 = a) \sim \mathcal{N}\left(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho(a - \mu_2)\,,\, (1-\rho^2)\sigma_1^2\right).$$

Problem 3: Building a state estimator

In this problem, we will incorporate measurements from the radar to improve our prediction of the landing time and location.

(a) Generate one realization of the complete trajectory of the projectile (including the effect of the noise $\boldsymbol{w}_t$). This trajectory describes the position and velocity of the projectile from its initial state until the time it lands. Use an initial condition sampled from the same Gaussian given before Problem 1. This is now the *ground truth* for the projectile's trajectory.

(b) Use the ground truth obtained in part (a) to generate a set of radar observations $\{\boldsymbol{y}_t\}_{t=1}^{T}$ for the entire trajectory.

(c) Run a Kalman filter for the first 10 seconds (until $t = 10$) to estimate the position and velocity of the projectile (mean and covariance). How does this prediction of the position compare to the case in Problem 1 part (e), where we did not use any radar measurements?

(d) Suppose the radar you were using stopped working after 10 seconds (i.e., after $t = 10$). Use all available information until $t = 10$ to predict the landing times and locations of the projectile. (*Hint: Use* $\widehat{\boldsymbol{x}}_{10|10}$ *and* $P_{10|10}$ *obtained from part (c) to predict future states and covariances using Equations* (5) *and* (6).)

(e) How does the prediction obtained in part (c) compare to that obtained in Problem 2, where no information other than the initial conditions was used?

Problem 4: Zooming out

Well done; you have reached the end of the project (and of the course)!

Before we bid each other good-bye for the summer, here are a few broader open-ended questions regarding the projectile tracking exercise. Remember, we are hoping to pique your curiosity, encourage you to think about the physical implications of this exercise, and tie in general concepts from estimation theory. While we are only looking for qualitative answers below, you may feel free to supplement your reasoning through simulations or even mathematical arguments!

(a) How would the predicted distributions of the landing time and of the landing location change if the radar worked for a longer duration of time?

(b) If the radar never stopped working, would we be able to observe the exact landing time and location?

(c) How would our prediction of the landing time and location change if the radar observations only involved the height of the projectile? In other words, how would the predictions change if

$$\boldsymbol{y}_t = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \boldsymbol{x}_t + \boldsymbol{r}_t \ ?$$

(d) What happens when measurement noise is very small or very large? Under which scenario would the prediction of the landing time and location be affected significantly if the radar stops working?

(e) How could you go about solving Problem 2 using Monte Carlo simulations?

(f) If you did not know anything about Kalman filters (i.e., the equations for prediction and update steps), would you have been able to solve Problem 3 using Monte Carlo simulations of the state?

(g) How important is the choice of Δt? What are the advantages and disadvantages of having a smaller Δt?

(h) In practice, how do you think we could estimate the covariance matrices Q and R?