

**Graduate Artificial Intelligence 15-780**

# Homework 1: *Probabilistic Inference*



Out on January 30  
Due on February 13

**Problem 1: Independence relations in probabilistic graphical models [Felipe, 40pts]**

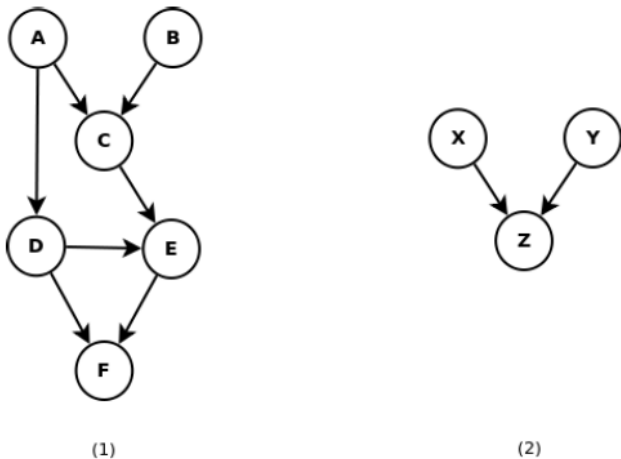


Figure 1: Directed models for Problem 1

a) For the independence relations below, say if they hold or not in Figure 1.1 and explain your answer. [8pts]

- (a)  $A \perp F | D, E$
- (b)  $B \perp D$
- (c)  $B \perp E | C$
- (d)  $B \perp F | E$

b) Construct the Undirected Model structure that captures all the following independence relations: [8 pts]

- $X_1 \perp X_2 | Y_1, Y_2$
- $Y_1 \perp Y_2 | X_1, X_2$
- $X_1 \not\perp X_2 | Y_1$
- $X_1 \not\perp X_2 | Y_2$
- $Y_1 \not\perp Y_2 | X_1$
- $Y_1 \not\perp Y_2 | X_2$

c) Is it possible to build a Directed Model such that its structure captures the same independence relations as in item b)? Show the structure or proof that it is not possible. [16 pts]

d) The Directed Model structure in Figure 1.2, known as v-structure, is such that  $X \perp Y$  and  $X \not\perp Y | Z$ . Considering  $X, Y$  and  $Z$  as binary random variables,  $P(X = 1) = 0.7$  and  $P(Y = 1) = 0.1$ , show a conditional probability table for  $P(Z | X, Y)$  such that  $X \perp Y | Z$  holds for this Directed Model. [8 pts]

## Problem 2: Broken Sparrow [John, 60pts]<sup>1</sup>

### The Story So Far...

Class, the unthinkable has happen.

An unmanned aerial vehicle (UAV) loaded with literally tens of hundreds of dollars worth of atmospheric sensors has gone rogue. Well, technically it malfunctioned. A standard flight started to go wrong at time-step  $t = 0$ . This was the last time-step that we received telemetry directly from the UAV.

Since then we have been tracking it from two far-off radar towers. At  $t = 0$  the UAV reported a system failure that not only disrupted its communication systems, but also knocked-out its guidance system.

You, and a special case of the Hidden Markov Model (HMM), are our only hope to find where the UAV is now.

### Technical Briefing: Linear Dynamic Systems

A *Linear Dynamic System* is a special case of the HMM where the latent (unobserved or hidden) state at  $t + 1$  and the observation at  $t$  are both linear functions of the latent state at  $t$  with additive Gaussian noise. This special case is extremely attractive mathematically because updating the model in the face of evidence can be written as a series of linear algebra operations.

Formally the system is defined by the following equations:

$$z_{t+1} = A \cdot z_t + w_t \quad (1)$$

$$x_t = C \cdot z_t + v_t. \quad (2)$$

Here  $z_t = (d_t^x, d_t^y, v_t^x, v_t^y, a_t^y, a_t^x)$  is the latent state and  $x_t = (a_t^x, a_t^y, b_t^x, b_t^y)$  is observed state (the observations of  $(x, y)$  from the two towers).  $A$  and  $C$  are matrices, called the *transition* and *observation* matrices. The noise terms,  $w_t$  and  $v_t$  are distributed by zero-mean multivariate Gaussians:

$$w_t \sim \mathcal{N}(w_t | \mathbf{0}, \Gamma) \quad (3)$$

$$v_t \sim \mathcal{N}(v_t | \mathbf{0}, \Sigma) \quad (4)$$

The initial state is also a multivariate Gaussian:

$$z_0 \sim \mathcal{N}(z_0 | \mu_0, V_0). \quad (5)$$

### Filling in the Numbers

- **Dynamics:**

We can model the transition of the UAV with the following simple 2D dynamics:

$$a_{t+1}^x = a_t^x \quad (6)$$

$$v_{t+1}^x = v_t^x + a_t^x \cdot \Delta t \quad (7)$$

$$d_{t+1}^x = d_t^x + v_t^x \cdot \Delta t + \frac{1}{2} a_t^x \cdot (\Delta t)^2 \quad (8)$$

<sup>1</sup>Modified from a problem suggested by Erik Zawadzki.

A similar set of equations could be written for the  $y$ -dimension. These equations define your transition matrix. We receive observations from the radar towers every second, so  $\Delta t = 1$ .

The UAV is reasonably aerodynamically stable, but because of its malfunctioning guidance system is believed to be accelerating randomly. In particular:

$$\Gamma = \begin{pmatrix} 0.0001 & & & & & \\ & 0.0001 & & & & \\ & & 0.0001 & & & \\ & & & 0.0001 & & \\ & & & & 0.1 & \\ & & & & & 0.1 \end{pmatrix} \quad (9)$$

#### • Observation

The two towers both provide unbiased but imprecise estimates of the UAV's  $(d_i^x, d_i^y)$  position every time-step. Luckily, their noise is completely independent of each other. Therefore, the system's observation noise has the following measurement-noise covariance matrix:

$$\Sigma_1 = \Sigma_2 = \begin{pmatrix} 100 & 10 \\ 10 & 100 \end{pmatrix} \quad (10)$$

$$\Sigma = \begin{pmatrix} \Sigma_1 & \\ & \Sigma_2 \end{pmatrix} \quad (11)$$

#### • Initial Position

The UAV was initially at  $(0, 0)m$  and was traveling at  $(6, 6)m/s$ . It was not accelerating at that time. However the initial telemetry data is also subject to noise;  $V_{-1} = \mathbb{I}_6$  (the  $6 \times 6$  identity matrix).

## Tasks

Your mission, should you choose to accept it:

- Draw the Bayes net for time steps  $t-1$ ,  $t$ , and  $t+1$  for arbitrary time-step  $t$ . *I.e.* draw all the variables for these three time steps, and all the edges that connect them. [10 pts]
- Understand how to infer the posterior distribution of latent state at  $t$  from the observed data at steps  $0, \dots, t$ . This inference in an LDS is performed by a *Kalman Filter*. More information can be found on Wikipedia's excellent *Kalman Filter* page, in §15.4 of Russell and Norvig, and in §13.3 of *Pattern Recognition and Machine Learning* by Bishop. [15 pts]

Comprehension questions for this task:

- Derive the expression for the joint distribution of two successive states and one observation:  $P(x_t, z_t, z_{t+1})$ . In this expression, assume that the observations prior to time  $t$  lead you to believe that the distribution of the previous state is  $z_t \sim \mathcal{N}(z_t | \mu_t, V_t)$ . Your answer should be the product multivariate Gaussians.
- Suppose that the distribution of state  $z_t \sim \mathcal{N}(z_t | \mu_t, V_t)$ . The distribution of  $P(z_{t+1} | x_{t+1})$  is normal. Please provide expressions for the mean and covariance. You do not have to provide the derivation for this answer, but cite your source if you do not.
- Plot the sensor data and sequence of posterior distribution means for the  $(x, y)$ -positions of the UAV. Figure 2 is an example of the kind of plot that we are looking for. (The actual data will be entirely different.) [25 pts]

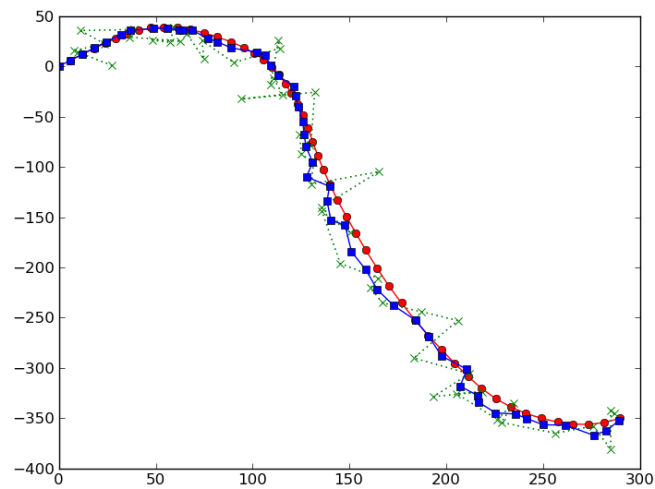


Figure 2: The blue line is the expected path, the green paths are the observed path from the two towers given in the files, and the red path is the true trajectory of the UAV. You are not given the data for the red path.

- d) Report UAV's posterior position distribution at  $t = 100$ . This is a multivariate Gaussian, and can be reported as its mean and  $6 \times 6$  covariance matrix. [10 pts]

**Note:** please pick an appropriate language to do this question in. Your language should be able to easily support linear algebra. If you want to build a test simulator, you also need to pick a language that allows sampling from multivariate Gaussians. MATLAB and Python/Numpy are both good choices.