An EM Algorithm for Learning in Controlled Linear Dynamical Systems

Ömer Deniz Akyıldız
deniz.akyildiz [at] boun.edu.tr
Department of Electrical Engineering
Bogaziçi University

2013

Abstract

In this report, we derive an expectation-maximization (EM) algorithm for learning in controlled dynamical systems. This report should be thought of as a simple extension of the algorithm presented in [1]. But in our case, we assume that the physical models of the objects and observation models are known, however the 'policies' of objects are not known. Starting from this point, we derive an EM algorithm for learning policies from data.

1 Introduction

A classical work in parameter estimation in linear dynamical systems is [1]. In [1], authors derive an EM algorithm to learn the parameters of a linear dynamical system. They consider the following dynamic model,

$$x_t = Ax_{t-1} + w_{t-1} \tag{1}$$

$$y_t = Cx_t + n_t \tag{2}$$

where $(x_t)_{t\geq 0}$ is a hidden stochastic process and $(y_t)_{t\geq 0}$ is the observation process and w_t and n_t are white noise sequences. They derive update rules (for EM) for A, C, Q, R, where Q and R noise covariances.

In this report, we proceed in a similar fashion (as an exercise). However, we assume our dynamical system is controlled [2]. A controlled linear dynamical system (CLDS) can be written as,

$$x_t = Ax_{t-1} + BLx_{t-1} + w_{t-1} \tag{3}$$
$$y_t = Cx_t + n_t \tag{4}$$

In general, control inputs are denoted by u_t. In our CLDS, we denote them with Lx_{t-1}. In theory, L is determined by minimizing a quadratic cost function[1]

$$\mathbb{E}\left\{ x_N^T Q_N x_N + \sum_{k=0}^{N-1} (x_k^T Q_k x_k + u_k^T R_k u_k) \right\} \tag{5}$$

For the derivation of L by minimizing this cost function, see [2].

An idea is to learn this cost function from observations is called as inverse optimal control or inverse reinforcement learning. However, from an EM viewpoint, deriving an EM algorithm to learn cost this function is pretty hard due to the recursive structure of the Algebraic Riccati Equation [2]. Rather, focusing on only to L matrix an EM algorithm can be devised. We shall assume A, B, C is known, but even if they are unknown, one can also derive an EM algorithm following the work [1]. The assumption about A, B, C makes sense because in general the physical models of objects are approximately known. For instance, if we want to track a vehicle, we know how its physical location evolve. However, we do not know if the driver accelerating or decelerating the vehicle. This is what we call as a policy.

In this work, we try to estimate policy matrix L. For a detailed discussion and derivation of policy matrix, see [2]. As we mentioned, we assume A, B, C is known, and further assume everything else is known including initial conditions etc. to simplify derivation – since the update rules for those quantities are known. Note that, to assume the dynamical system is controlled, we need the pair (A, B) should be controllable and the pair (A, C) should be observable [2]. We assume these assumptions hold.

2 The EM algorithm

EM algorithm aims to maximize the log likelihood of complete data. For a treatment of EM algorithm, see [3]. First of all, we write the likelihood as follows,

$$p(x_{1:T}, y_{1:T}|L) = p(x_1) \prod_{t=2}^{T} p(x_t|x_{t-1}, L) \prod_{t=1}^{T} p(y_t|x_t) \tag{6}$$

[1]Do not confuse: These Q and R are not noise covariances. They characterise the interaction over state variables.

by using the conditional independence assumptions implicit in our model. Our model is an instance of hidden Markov models and these models factorizes over observation and transition distributions [4]. By departing this point, we write the log likelihood as follows.

$$\log p(x_{1:T}, y_{1:T}|L) = \log p(x_1) + \sum_{t=2}^{T} \log p(x_t|x_{t-1}, L) + \sum_{t=1}^{T} \log p(y_t|x_t) \quad (7)$$

In a typical expectation maximization setup, we aim to perform two step: (1) E-step, (2) M-step. In E-step, we aim to calculate following expectation,

$$\langle \log p(x_{1:T}, y_{1:T}|L) \rangle_{p(x_{1:T}|y_{1:T}, L^*)} \quad (8)$$

where L^* is a fixed parameter. Then, in maximization step, we aim to update this L^* by setting derivative of this expectation with respect to L to zero. We start with the maximization step (M-step) and denote expectations in the equations. Then, we show E-step, i.e., how to compute these expectations.

2.1 M-step

The important term for us is $p(x_t|x_{t-1}, L)$. Let us write it explicitly,

$$p(x_t|x_{t-1}, L) = \mathcal{N}(x_t; (A + BL)x_{t-1}, Q) \quad (9)$$

Suppose A, B, Q is known. More explicitly,

$$p(x_t|x_{t-1}) = \exp\left\{ -\frac{1}{2} \left(x_t - (A + BL)x_{t-1}\right)^\top Q^{-1} \left(x_t - (A + BL)x_{t-1}\right) \right\} (2\pi)^{-k/2} |Q|^{-1/2} \quad (10)$$

If we use this formula and try to devise a formula for update rule L, we have to minimize following cost function[2] [3] (will be extended):

$$\mathcal{Q}(\theta, \theta^{\text{old}}) = -\frac{T-1}{2} \log |Q| - \mathbb{E}_{x_{1:T}|\theta^{\text{old}}} \left[\frac{1}{2} \sum_{t=2}^{T} (x_t - (A + BL)x_{t-1})^\top Q^{-1} (x_t - (A + BL)x_{t-1}) \right] \quad (11)$$

First, consider the term $(x_t - (A + BL)x_{t-1})^\top Q^{-1}(x_t - (A + BL)x_{t-1})$. We can expand this term as follows.

$$\begin{aligned} x_t^\top Q^{-1} x_t - x_t^\top Q^{-1}(A + BL)x_{t-1} - ((A + BL)x_{t-1})^\top Q^{-1} x_t \\ + ((A + BL)x_{t-1})^\top Q^{-1}((A + BL)x_{t-1}) \end{aligned}$$

[2] $(\cdot)^\top$ denotes transpose.

By continuing the derivation,

$$\begin{aligned} x_t^\top Q^{-1}x_t - x_t^\top Q^{-1}(A+BL)x_{t-1} - x_{t-1}^\top (A+BL)^\top Q^{-1}x_t \\ +x_{t-1}^\top (A+BL)^\top Q^{-1}(A+BL)x_{t-1} \end{aligned}$$

So, this is equal to the following (we are expanding the last term),

$$\begin{aligned} x_t^\top Q^{-1}x_t - x_t^\top Q^{-1}(A+BL)x_{t-1} - x_{t-1}^\top (A+BL)^\top Q^{-1}x_t \\ +x_{t-1}^\top A^\top Q^{-1}(A+BL)x_{t-1} + x_{t-1}^\top (BL)^\top Q^{-1}(A+BL)x_{t-1} \end{aligned}$$

One more step,

$$\begin{aligned} x_t^\top Q^{-1}x_t - x_t^\top Q^{-1}(A+BL)x_{t-1} - x_{t-1}^\top (A+BL)^\top Q^{-1}x_t + x_{t-1}^\top A^\top Q^{-1}Ax_{t-1} \\ +x_{t-1}^\top A^\top Q^{-1}BLx_{t-1} + x_{t-1}^\top (BL)^\top Q^{-1}Ax_{t-1} + x_{t-1}^\top (BL)^\top Q^{-1}BLx_{t-1} \end{aligned}$$

Hence, by expanding all terms, we can write this expression as,

$$\begin{aligned} x_t^\top Q^{-1}x_t - x_t^\top Q^{-1}Ax_{t-1} - x_t^\top Q^{-1}BLx_{t-1} - x_{t-1}^\top A^\top Q^{-1}x_t - x_{t-1}^\top (BL)^\top Q^{-1}x_t + \\ x_{t-1}^\top A^\top Q^{-1}Ax_{t-1} + x_{t-1}^\top A^\top Q^{-1}BLx_{t-1} + x_{t-1}^\top (BL)^\top Q^{-1}Ax_{t-1} + x_{t-1}^\top (BL)^\top Q^{-1}BLx_{t-1} \end{aligned}$$

We can write it as (by using the properties of trace [5])[3],

$$\begin{aligned} F_t = \mathrm{Tr}\Big[& Q^{-1}x_t x_t^\top - A^\top Q^{-1}x_t x_{t-1}^\top - (BL)^\top Q^{-1}x_t x_{t-1}^\top - A^\top Q^{-1}x_t x_{t-1}^\top - (BL)^\top Q^{-1}x_t x_{t-1}^\top \\ & +A^\top Q^{-1}Ax_{t-1}x_{t-1}^\top + A^\top Q^{-1}BLx_{t-1}x_{t-1}^\top + (BL)^\top Q^{-1}Ax_{t-1}x_{t-1}^\top + (BL)^\top Q^{-1}BLx_{t-1}x_{t-1}^\top \Big] \end{aligned}$$

We call this quantity as F_t to use in below equations. Now, if we compute,

$$\frac{\partial \mathcal{Q}(\theta,\theta^{\text{old}})}{\partial L} = -\frac{1}{2}\mathbb{E}_{x_{1:T}|\theta^{\text{old}}}\left[\sum_{t=2}^{T} \frac{\partial (x_t-(A+BL)x_{t-1})^\top Q^{-1}(x_t-(A+BL)x_{t-1})}{\partial L}\right] \tag{12}$$

If we use the properties derived in above equations, it is equal to

$$\frac{\partial \mathcal{Q}(\theta,\theta^{\text{old}})}{\partial L} = -\frac{1}{2}\mathbb{E}_{x_{1:T}|\theta^{\text{old}}}\left[\sum_{t=2}^{T} \frac{\partial F_t}{\partial L}\right] \tag{13}$$

[3] Since $Q^{-1} = {Q^{-1}}^\top$

We can expand the derivatives,

$$\frac{\partial \mathcal{Q}(\theta,\theta^{\text{old}})}{\partial L} = -\frac{1}{2}\mathbb{E}_{x_{1:T}|\theta^{\text{old}}}\left[\sum_{t=2}^{T}\left(-\frac{\partial \text{Tr}\left[(BL)^\top Q^{-1}x_t x_{t-1}^\top\right]}{\partial L} - \frac{\partial \text{Tr}\left[(BL)^\top Q^{-1}x_t x_{t-1}^\top\right]}{\partial L}\right.\right.$$
$$\left.\left.+\frac{\partial \text{Tr}\left[A^\top Q^{-1}BLx_{t-1}x_{t-1}^\top\right]}{\partial L} + \frac{\partial \text{Tr}\left[(BL)^\top Q^{-1}Ax_{t-1}x_{t-1}^\top\right]}{\partial L} + \frac{\partial \text{Tr}\left[(BL)^\top Q^{-1}BLx_{t-1}x_{t-1}^\top\right]}{\partial L}\right)\right] \tag{14}$$

More painful algebra (put $L^\top B^\top$ instead of $(BL)^\top$,

$$\frac{\partial \mathcal{Q}(\theta,\theta^{\text{old}})}{\partial L} = -\frac{1}{2}\mathbb{E}_{x_{1:T}|\theta^{\text{old}}}\left[\sum_{t=2}^{T}\left(-\frac{\partial \text{Tr}\left[L^\top B^\top Q^{-1}x_t x_{t-1}^\top\right]}{\partial L} - \frac{\partial \text{Tr}\left[L^\top B^\top Q^{-1}x_t x_{t-1}^\top\right]}{\partial L}\right.\right.$$
$$\left.\left.+\frac{\partial \text{Tr}\left[A^\top Q^{-1}BLx_{t-1}x_{t-1}^\top\right]}{\partial L} + \frac{\partial \text{Tr}\left[L^\top B^\top Q^{-1}Ax_{t-1}x_{t-1}^\top\right]}{\partial L} + \frac{\partial \text{Tr}\left[L^\top B^\top Q^{-1}BLx_{t-1}x_{t-1}^\top\right]}{\partial L}\right)\right] \tag{15}$$

Since we can change the order in the trace operator, we can make this expression more differentiation-friendly,

$$\frac{\partial \mathcal{Q}(\theta,\theta^{\text{old}})}{\partial L} = -\frac{1}{2}\mathbb{E}_{x_{1:T}|\theta^{\text{old}}}\left[\sum_{t=2}^{T}\left(-\frac{\partial \text{Tr}\left[B^\top Q^{-1}x_t x_{t-1}^\top L^\top\right]}{\partial L} - \frac{\partial \text{Tr}\left[B^\top Q^{-1}x_t x_{t-1}^\top L^\top\right]}{\partial L}\right.\right.$$
$$\left.\left.+\frac{\partial \text{Tr}\left[x_{t-1}x_{t-1}^\top A^\top Q^{-1}BL\right]}{\partial L} + \frac{\partial \text{Tr}\left[B^\top Q^{-1}Ax_{t-1}x_{t-1}^\top L^\top\right]}{\partial L} + \frac{\partial \text{Tr}\left[B^\top Q^{-1}BLx_{t-1}x_{t-1}^\top L^\top\right]}{\partial L}\right)\right] \tag{16}$$

Now we can take derivatives[4],

$$\begin{aligned}\frac{\partial \mathcal{Q}(\theta,\theta^{\text{old}})}{\partial L} = -\frac{1}{2}\mathbb{E}_{x_{1:T}|\theta^{\text{old}}}\Bigg[\sum_{t=2}^{T}\big(&-\left[B^\top Q^{-1}x_t x_{t-1}^\top\right] - \left[B^\top Q^{-1}x_t x_{t-1}^\top\right]\\ &+\left[B^\top Q^{-1}Ax_{t-1}x_{t-1}^\top\right] + \left[B^\top Q^{-1}Ax_{t-1}x_{t-1}^\top\right]\\ &+\left[B^\top Q^{-1}BLx_{t-1}x_{t-1}^\top + B^\top Q^{-1}BLx_{t-1}x_{t-1}^\top\right]\big)\Bigg]\end{aligned} \tag{17}$$

Our aim is to marginalize L by setting,

$$\frac{\partial \mathcal{Q}(\theta,\theta^{\text{old}})}{\partial L} = 0$$

[4]From [5], we have $\frac{\partial Tr[MXNX^\top]}{\partial X} = M^\top XN^\top + MXN$. Here we have $M = B^\top Q^{-1}B$ and $N = x_{t-1}x_{t-1}^T$. Luckily, $M = M^T$ and $N = N^T$.

For the sake of simplicity, let us denote the matrices as we denote expectations,

$$E_{t|t-1} = \mathbb{E}_{x_{1:T}|\theta^{\text{old}}}\left[x_t x_{t-1}^\top\right] \tag{18}$$
$$E_{t-1|t-1} = \mathbb{E}_{x_{1:T}|\theta^{\text{old}}}\left[x_{t-1} x_{t-1}^\top\right] \tag{19}$$

To simplify more, we adapt a new notation,

$$E^s_{t|t-1} = \sum_{t=2}^{T} E_{t|t-1} \tag{20}$$
$$E^s_{t-1|t-1} = \sum_{t=2}^{T} E_{t-1|t-1} \tag{21}$$
$$\tag{22}$$

where s stands for summation from 2 to T. Then, if we set the derivative to zero, we have,

$$\begin{aligned} &B^\top Q^{-1} E^s_{t|t-1} + B^\top Q^{-1} E^s_{t|t-1} - B^\top Q^{-1} A E^s_{t-1|t-1} - B^\top Q^{-1} A E^s_{t-1|t-1} \\ &= \left(B^\top Q^{-1} B L E^s_{t-1|t-1} + B^\top Q^{-1} B L E^s_{t-1|t-1}\right) \end{aligned} \tag{23}$$

So let us denote $B^\top Q^{-1} B$ as M. And denote $E^s_{t-1|t-1}$ as N. Denote $Z = B^\top Q^{-1} E^s_{t|t-1} - B^\top Q^{-1} A E^s_{t-1|t-1}$. Hence our problem is finding an expression for L by using the following equation,

$$Z = MLN \tag{24}$$

Then, we have,

$$L = M^{-1} Z N^{-1} \tag{25}$$

2.2 E-step

For a given estimate of L, one can use the exact same E-step given in [1]. One has to perform a forward-filtering, backward-smoothing algorithm for linear dynamical systems, i.e., the Kalman filtering followed by the RTS smoothing to estimate the expectations given in the M-step. One can consult to [1] to employ this algorithm, only by modifying the A matrix with $A + B\hat{L}$ where $\hat{L}$ is the current estimate of the matrix L.

References

[1] Zoubin Ghahramani and Geoffrey E. Hinton. Parameter estimation for linear dynamical systems. Technical report, 1996.

[2] D. P. Bertsekas. *Dynamic Programming and Optimal Control, Vol. I, 2nd Ed.* Athena Scientific, Belmont, MA, 2001.

[3] Christopher M. Bishop. *Pattern Recognition and Machine Learning.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[4] Olivier Cappé, Eric Moulines, and Tobias Ryden. *Inference in Hidden Markov Models.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[5] K. B. Petersen and M. S. Pedersen. The matrix cookbook, nov 2012. Version 20121115.