# Kalman filter and smoother

Contents

The Kalman filter is a method of estimating the current state of a dynamical system, given the observations so far. The underlying model is a hidden Markov model (HMM) in which everything is multivariate normal—so in particular, the hidden variables are continuous, rather than discrete. The Kalman filter is actually just the forward algorithm, except that each step can be computed analytically due to the magic of Gaussians. As one might expect, there is also a backward algorithm (or something very similar), and this is referred to as the Kalman smoother. The Kalman smoother allows one to refine estimates of previous states, in the light of later observations. As in the case of discrete-space HMMs, the results of the Kalman filter and smoother can also be combined with expectation-maximization to estimate the parameters of the model. I think it is fair to say that the Kalman filter is one of the most important algorithms of the 20th century.

1 Background

- Around 1960, the United States and the Soviet Union were developing rocket technology for their space programs (and, no coincidence, for intercontinental ballistic missiles as well).

- At the same time, Rudolph Kalman in the US and Ruslan Stratonovich in the USSR were developing methods for efficiently and accurately estimating the state of a dynamical system by accumulating noisy measurements from many different instruments over time. Kalman's method would later become known as the Kalman filter, and is a special case of Stratonovich's method. Early contributions were also made by Thorvald Thiele, Peter Swerling, and Richard Bucy.

- When Kalman visited NASA Ames Research Center, Stanley Schmidt realized the applicability of Kalman's work to the problem of navigation and control of aircraft and spacecraft, and the Kalman filter became an integral part of the Apollo navigation computer.

- Today, essentially all high-performance navigation systems use a Kalman filter or some variant thereof. The method is used in rockets, missiles, spacecraft including the international space station, unmanned aerial vehicles, ground robots, and recently, self-driving cars.

2 Model

- The Kalman filter and smoother are based on the following probabilistic model.

- As in a discrete-space HMM, the sequence of observed data $x_1, \ldots, x_n$ is modeled jointly along with a sequence of hidden states $z_1, \ldots, z_n$ by a distribution that respects the graph:
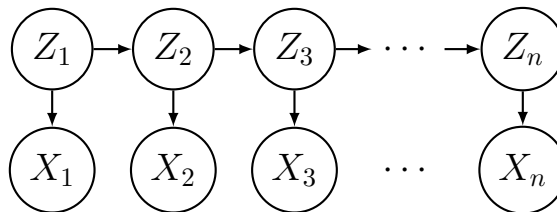
In other words, we assume

$$p(x_{1:n}, z_{1:n}) = p(z_1)p(x_1|z_1) \prod_{j=2}^{n} p(z_j|z_{j-1})p(x_j|z_j).$$

- However, unlike a discrete-space HMM, each hidden state z_j is modeled as a continuous random variable in $\mathbb{R}^d$, following a multivariate normal distribution.

- Specifically, the initial distribution $p(z_1)$, the transition distributions $p(z_j|z_{j-1})$ (a.k.a. the "process model"), and the emission distributions $p(x_j|z_j)$ (a.k.a. the "measurement model") are assumed to be

$$p(z_1) = \mathcal{N}(z_1 \mid \mu_0, V_0)$$
$$p(z_j|z_{j-1}) = \mathcal{N}(z_j \mid Fz_{j-1}, Q)$$
$$p(x_j|z_j) = \mathcal{N}(x_j \mid Hz_j, R)$$

  where

  - $z_j \in \mathbb{R}^d$ (the state of the system at time step j),
  - $x_j \in \mathbb{R}^D$ (the measurements at time step j),
  - $\mu_0 \in \mathbb{R}^d$ is an arbitrary vector (the initial mean, our "best guess" at the initial state),
  - $V_0 \in \mathbb{R}^{d \times d}$ is a symmetric positive definite matrix (the initial covariance matrix, quantifying our uncertainty about the initial state),
  - $F \in \mathbb{R}^{d \times d}$ is an arbitrary matrix (capturing the physics of the process, or a linear approximation thereof),
  - $Q \in \mathbb{R}^{d \times d}$ is a symmetric positive definite matrix (quantifying the noise/error in the process),
  - $H \in \mathbb{R}^{D \times d}$ is an arbitrary matrix (relating the measurements to the state),
  - $R \in \mathbb{R}^{D \times D}$ is a symmetric positive definite matrix (quantifying the noise/error in the measurements).