Lecture 30

Filtering, Smoothing, & Forecasting

Arnab Hazra

Things to cover (recap)

- We will cover the concepts of
  - prediction
  - filtering
  - smoothing

  state space models and include their derivations.

Definitions

- A primary aim of state space models is to produce estimators for the underlying unobserved signal $\boldsymbol{X}_t$, given the data $\mathcal{Y}_{1:s} = \{\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_s\}$.
- State estimation is an essential component of parameter estimation.
- In addition to these estimates, we would also want to measure their precision.
- When $s < t$, the problem is called forecasting or prediction.
- When $s = t$, the problem is called filtering.
- When $s > t$, the problem is called smoothing.

DLM with covariates (recap)

- The ARMAX model involves covariates that may enter into the states or into the observations.

- In this case, we suppose we have an $r \times 1$ vector of inputs $\boldsymbol{u}_t$, and write the model as

$$\boldsymbol{X}_t = \Phi \boldsymbol{X}_{t-1} + \gamma \boldsymbol{u}_t + \boldsymbol{W}_t$$

$$\boldsymbol{Y}_t = \boldsymbol{A}_t \boldsymbol{X}_t + \Gamma \boldsymbol{u}_t + \boldsymbol{V}_t$$

- Here γ is $p \times r$ and Γ is $q \times r$; either of these matrices may be the zero matrix.

Notations

We will use the following definitions:

- $$\boldsymbol{X}_t^s = E(\boldsymbol{X}_t | \mathcal{Y}_{1:s})$$
- $$\boldsymbol{P}_{t_1,t_2}^s = E[(\boldsymbol{X}_{t_1} - \boldsymbol{X}_{t_1}^s)(\boldsymbol{X}_{t_2} - \boldsymbol{X}_{t_2}^s)']$$
- $$\boldsymbol{P}_t^s = \boldsymbol{P}_{t,t}^s = \mathrm{Cov}(\boldsymbol{X}_t - \boldsymbol{X}_t^s)$$
- Due to Gaussian assumption,

  $$\boldsymbol{P}_{t_1,t_2}^s = E[(\boldsymbol{X}_{t_1} - \boldsymbol{X}_{t_1}^s)(\boldsymbol{X}_{t_2} - \boldsymbol{X}_{t_2}^s)' | \mathcal{Y}_{1:s}],$$

  follows from the fact that the covariance matrix between $(\boldsymbol{X}_t - \boldsymbol{X}_t^s)$ and $\mathcal{Y}_{1:s}$, for any t and s, is zero.

The Kalman Filter

- With initial conditions $\boldsymbol{X}_0^0 = \mu_0$ and $\boldsymbol{P}_0^0 = \Sigma_0$, for $t = 1, \dots, T$,

$$\boldsymbol{X}_t^{t-1} = \Phi \boldsymbol{X}_{t-1}^{t-1} + \gamma \boldsymbol{u}_t, \quad \boldsymbol{P}_t^{t-1} = \Phi \boldsymbol{P}_{t-1}^{t-1} \Phi' + \boldsymbol{Q}$$

  with

$$\boldsymbol{X}_t^t = \boldsymbol{X}_t^{t-1} + \boldsymbol{K}_t(\boldsymbol{Y}_t - \boldsymbol{A}_t \boldsymbol{X}_t^{t-1} - \Gamma \boldsymbol{u}_t), \quad \boldsymbol{P}_t^t = [\boldsymbol{I} - \boldsymbol{K}_t \boldsymbol{A}_t] \boldsymbol{P}_t^{t-1}.$$

- Here the Kalman gain is

$$\boldsymbol{K}_t = \boldsymbol{P}_t^{t-1} \boldsymbol{A}_t' [\boldsymbol{A}_t \boldsymbol{P}_t^{t-1} \boldsymbol{A}_t' + \boldsymbol{R}]^{-1}.$$

- Prediction for $t > T$ is accomplished via $\boldsymbol{X}_t^{t-1} = \Phi \boldsymbol{X}_{t-1}^{t-1} + \gamma \boldsymbol{u}_t$ and $\boldsymbol{P}_t^{t-1} = \Phi \boldsymbol{P}_{t-1}^{t-1} \Phi' + \boldsymbol{Q}$ with initial conditions $\boldsymbol{X}_T^T$ and $\boldsymbol{P}_T^T$.

Proofs on board

The Kalman Filter (contd.)

- Important byproducts of the filter are the innovations (prediction errors)

$$\varepsilon_t = \boldsymbol{Y}_t - E(\boldsymbol{Y}_t|\mathcal{Y}_{1:(t-1)}) = \boldsymbol{Y}_t - \boldsymbol{A}_t X_t^{t-1} - \Gamma \boldsymbol{u}_t,$$

  and the corresponding variance-covariance matrices

$$\Sigma_t \stackrel{def}{=} \mathrm{Cov}(\varepsilon_t) = \mathrm{Cov}[\boldsymbol{A}_t(\boldsymbol{X}_t - \boldsymbol{X}_t^{t-1}) + \boldsymbol{V}_t] = \boldsymbol{A}_t \boldsymbol{P}_t^{t-1} \boldsymbol{A}_t' + \boldsymbol{R}$$

  for $t = 1, \ldots, T$.

- We assume that $\Sigma_t > 0$ (is positive definite), which is guaranteed, for example, if $\boldsymbol{R} > 0$. This assumption is not necessary and may be relaxed.

Proofs on board

The Kalman Smoother

For the DLM with covariates, with initial conditions $\boldsymbol{X}_T^T$ and $\boldsymbol{P}_T^T$ obtained Kalman Filter, for $t = T, T-1, \ldots, 1$,

- $$\boldsymbol{X}_{t-1}^T = \boldsymbol{X}_{t-1}^{t-1} + \boldsymbol{J}_{t-1}(\boldsymbol{X}_t^T - \boldsymbol{X}_t^{t-1}),$$
- $$\boldsymbol{P}_{t-1}^T = \boldsymbol{P}_{t-1}^{t-1} + \boldsymbol{J}_{t-1}(\boldsymbol{P}_t^T - \boldsymbol{P}_t^{t-1})\boldsymbol{J}_{t-1}',$$

  where
  $$\boldsymbol{J}_{t-1} = \boldsymbol{P}_{t-1}^{t-1}\Phi'[\boldsymbol{P}_t^{t-1}]^{-1}.$$

Proofs on board

The Lag-One Covariance Smoother

For the DLM with covariates, with $\boldsymbol{K}_t$, $\boldsymbol{J}_t$ $(t = 1, \ldots, T)$, and $\boldsymbol{P}_T^T$ obtained from Kalman filter and Kalman smoother, and with initial condition

$$\boldsymbol{P}_{T,T-1}^T = (\boldsymbol{I}_q - \boldsymbol{K}_T \boldsymbol{A}_T)\Phi \boldsymbol{P}_{T-1}^{T-1},$$

- For $t = T, T-1, \ldots, 2$,

$$\boldsymbol{P}_{t-1,t-2}^T = \boldsymbol{P}_{t-1}^{t-1} \boldsymbol{J}_{t-2}' + \boldsymbol{J}_{t-1}(\boldsymbol{P}_{t,t-1}^T - \Phi \boldsymbol{P}_{t-1}^{t-1})\boldsymbol{J}_{t-2}'.$$

Proofs on board

Thank you!