Black-Litterman-Bayes, Kalman Filter, ICA

Kaiwen Zhou

Contents

1 Topic: Black-Litterman-Bayes

Problem 1.1. *This question refers to the article "On The Bayesian Interpretation Of Black-Litterman" (Kolm and Ritter, EJOR 2017).*

(a) *Derive formulas (10)-(11) using the properties of the multivariate normal distribution in the slides "Bayesian Modeling: Introduction".*

(b) ***(Extra credit)*** *Derive formulas (22)-(26) using the same properties.*

Solution:

(a) Derive formulas (10)-(11) using the properties of the multivariate normal distribution in the slides "Bayesian Modeling: Introduction".

(b) **(Extra credit)** Derive formulas (22)-(26) using the same properties.

2 Topic: Important Properties Of the Kalman Filter

Problem 2.1. *Choose three out of the four subproblems. It is optional to solve the other one for extra credit.*

(a) *In class we derived the Kalman filter under the assumption that* $\mathbb{E}\left[\mathbf{w}_t\mathbf{v}_t'\right] = \mathbf{0}$. *Now, derive the Kalman filter under the assumption that* $\mathbb{E}\left[\mathbf{w}_t\mathbf{v}_t'\right] = \mathbf{M}_t$.

(b) *Using the notation from class, let us denote the error between the true and estimated states by the random variable* $\tilde{\mathbf{x}}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t$. *We denote by* $\mathbf{W}_t$ *a known positive definite matrix. Show that the Kalman filter is the solution to the problem*

$$\min_{\hat{\mathbf{x}}_t} \mathbb{E}\left[\tilde{\mathbf{x}}_t'\mathbf{W}_t\tilde{\mathbf{x}}_t\right] \tag{1}$$

(c) *Now let us assume that* $\mathbf{w}_t$ *and* $\mathbf{v}_t$ *have zero mean, are uncorrelated with covariance matrices* $\mathbf{Q}_t$ *and* $\mathbf{R}_t$*, respectively (but they are no longer Gaussian). Show that the Kalman filter is the best linear solution to Equation (1). In other words, the Kalman filter is the best filter that is a linear combination of the measurements,* $\mathbf{y}_t$.

(d) *In class we derived the Kalman filter under the assumption that* $\mathbf{w}_t$ *and* $\mathbf{v}_t$ *are uncolored (i.e. each is serially uncorrelated).*

   (i) *Derive the Kalman filter under the assumption that* $\mathbf{w}_t$ *is a VAR(1) process with known system matrix.*

   (ii) *Derive the Kalman filter under the assumption that* $\mathbf{v}_t$ *is a* $\mathrm{VAR}(1)$ *process with known system matrix.*

   Hint: For (d), both (i) and (ii) can be solved by properly augmenting the state equations. Can you find a solution to (ii) where one does not have to augment the state? Is it possible to do so for (i) - why, or why not?

3 Topic: Kalman filter's Application on Financial Problems — Pairs Trading

Problem 3.1. *In this question we consider a basic pairs trading strategy between two stocks with prices* p_t^A *and* p_t^B *at time* t. *We denote the spread between them by* $s_t := \log\left(p_t^A\right) - \log\left(p_t^B\right)$ *and assume the spread follows an Ornstein-Uhlenbeck process*

$$ds_t = \kappa\left(\theta - s_t\right)dt + \sigma dB_t$$

where dB_t *is a standard Brownian motion. In other words, the spread reverts to its mean* $\theta \in \mathbb{R}$ *at the speed* $\kappa \in \mathbb{R}_+$*and volatility* $\sigma \in \mathbb{R}_+$.

(a) *Show that the discrete time solution of (2) is Markovian, that is*

$$s_k = \mathbb{E}\left[s_k \mid s_{k-1}\right] + \varepsilon_k$$

where $k = 1, 2, \ldots$*, and* ε_k *is a random process with zero mean and variance equal to* $\sigma_{\varepsilon,k}^2 = \mathbb{V}\left[s_k \mid s_{k-1}\right]$. *(Hint: You can derive the discrete solution explicitly.)*

(b) *Propose a methodology for updating the parameters* θ, κ *using the Kalman filter and describe how you would use it to trade the stock pair.*

(c) *Test your methodology from (b) on simulated data. In particular, (i) simulate (2) from known parameters* θ, κ *and* σ*, and then (ii) use the Kalman filter to recover them. You do not need to implement the Kalman filter from scratch; you are welcome to use a Kalman implementation from a Python package such as pykalman. How do you obtain a good estimate of* σ *?*

(d) *Repeat the same experiment from (c), but this time simulate (2) first with* κ *having the same value as above and then suddenly changing it to another value such that the half-life of the spread is* 50% *of its original value. How long does it take the Kalman filter to adjust? Can you make adjustment to your filter in order to speed up the time it takes the Kalman filter to adjust?*

Hint: For (c) and (d), think about how you are going to demonstrate the results using appropriate graphs, etc.

4 Topic: Kalman filter's Application on Financial Problems — Index Tracking Portfolios

Problem 4.1. *It is common in portfolio management to build so-called (index) tracking portfolios. Let us assume we are observing the return of the S&P 500 benchmark index, $r_{b,t}$. Now, let us pick a subset of 50 stocks from the constituents of this index. We will use these stocks to build a tracking portfolio for the index. For example, this could be the 50 companies in the index with the largest market cap. We denote the returns of these 50 stock by $\mathbf{r}_t \in \mathbb{R}^{50}$. The goal of finding a tracking portfolio is to find a dynamic trading strategy of the 50 stocks such that $\beta_t' \mathbf{r}_t \approx r_{b,t}$, where β_t denotes the holdings of the tracking portfolio.*

(a) In this part, we assume that the covariance matrix of returns of the stocks in the $\mathrm{S\&P500}$, $\boldsymbol{\Sigma}$, is given and constant through time. Find the portfolio of these 50 stocks that minimizes the tracking error to $r_{b,t}$, i.e. find the solution to

$$\boldsymbol{\beta}_t^* = \operatorname{argmin}_{\boldsymbol{\beta}_t} \sqrt{\mathbb{V}\left[r_{b,t} - \boldsymbol{\beta}_t' \mathbf{r}_t\right]}.$$

What specific property does β_t have here?

(b) In this part, we no longer assume that covariances amongst stocks are time invariant. Propose a solution to minimizing the tracking error using the Kalman filter.

(c) Download daily market data and create an example that illustrates your methodology. How does the tracking portfolio of the Kalman filter perform?

Hint: For (c), compare your Kalman filter to a solution based on (a) where the covariance matrix is estimated on rolling windows. Can you match the performance of the Kalman filter with the simpler methodology in (a) by appropriately choosing the length of the rolling window?

Topic: Another Interpretation for Independent Component Analysis (ICA)

Problem 4.2. *In class we considered the principally-independent component analysis method which essentially was the truncated rank- K SVD of a matrix X followed by an ICA rotation of the left singular components:*

$$\mathbf{X} \simeq \mathbf{U}\mathbf{S}\mathbf{V}' = \mathbf{U}_I \mathbf{S}_I \mathbf{V}_I$$

where $\mathbf{U}_I = \mathbf{U}\mathbf{A}_I$ with $\mathbf{A}_I = \underset{\mathbf{A}, \mathbf{A}'\mathbf{A} = I_K}{\operatorname{argmax}} \left|k_\ell(\mathbf{U}\mathbf{A})\right|$, $k_\ell(\mathbf{G})$ being any centered cumulant of order $\ell \geq 3$ which for all practical purposes can be considered a non-linear (activation)

function applied to each of the entries of $\mathbf{G}$. Furthermore the matrix $\mathbf{V}_I$ was defined as $\mathbf{V}_I' := \mathbf{D}^{-1}\mathbf{S}^{-1}\mathbf{A}_I \mathbf{S}\mathbf{V}'$ where $\mathbf{D}$ was chosen so that $\mathbf{V}_I$ has unimodular columns.

(a) Show that $\mathbf{D}$ is a diagonal matrix

(b) Show that $\mathbf{S}_I = \mathbf{S}\mathbf{D}$ is diagonal such that $\operatorname{Tr}\left(\mathbf{S}_I^2\right) = \operatorname{Tr}\left(\mathbf{S}^2\right)$.

(c) Show that the method can be derived as the limit, $\lambda^2 \to 0$, of the optimization

$$\mathbf{U}_I, \mathbf{S}_I, \mathbf{V}_I = \underset{\mathbf{P}, \mathbf{Q}: \mathbf{P}'\mathbf{P} = \operatorname{diag}(\mathbf{Q}'\mathbf{Q}) = \mathbf{I}_k}{\operatorname{argmin}} \left\|\mathbf{X} - \mathbf{P}\mathbf{R}\mathbf{Q}'\right\|_F - \lambda^2 \left|k_\ell(\mathbf{P})\right|.$$

(d) Show that an alternative objective function achieving the same result is

$$\mathbf{U}_I, \mathbf{S}_I, \mathbf{V}_I = \underset{\mathbf{P}, \mathbf{Q}: \mathbf{P}'\mathbf{P} = \operatorname{diag}(\mathbf{Q}'\mathbf{Q}) = \mathbf{I}_k}{\operatorname{argmin}} \left\|\mathbf{X} - \mathbf{P}\mathbf{R}\mathbf{Q}'\right\|_F - \lambda^2 J(\mathbf{P}),$$

where $J[\mathbf{x}] := H\left[\mathbf{x}_{gauss}\right] - H[\mathbf{x}]$ is the negentropy and $J(\mathbf{P})$ is the sum of the negentropies of all the columns of $\mathbf{P}$.