科技部補助專題研究計畫成果報告
期末報告

動態系統的貝氏分析(第2年)

計 畫 類 別 ： 個別型計畫
計 畫 編 號 ： MOST 102-2118-M-004-003-MY2
執 行 期 間 ： 103年08月01日至104年08月31日
執 行 單 位 ： 國立政治大學統計學系

計畫主持人： 翁久幸

計畫參與人員： 碩士班研究生-兼任助理人員：林冠謀
碩士班研究生-兼任助理人員：楊智博
碩士班研究生-兼任助理人員：陳弘叡

報 告 附 件 ： 出席國際會議研究心得報告及發表論文

處 理 方 式 ：
1.公開資訊：本計畫可公開查詢
2.「本研究」是否已有嚴重損及公共利益之發現：否
3.「本報告」是否建議提供政府單位施政參考：否

中 華 民 國 104年09月08日

中文摘要： 在過去數年，我們已經相當程度探索 Woodroofe-Stein 等式及其在序貫分析和貝式統計之應用。本計畫著重在動態模型之應用及其與卡爾曼 fileter 之關聯 。我們得到一種新的方法來推導卡爾曼增益(Kalman gain)。本計畫之副產品是將目前 Woodroofe-Stein 等式進一步推展， 得到多元的 Gram-Charlier 級數。

中文關鍵詞： 貝氏分析，動態系統，卡爾曼濾波，Woodroofe-Stein 等式

英文摘要： In the past years we have explored Woodroofe-Stein ‘s identity and its applications to sequential analysis and Bayesian statistics. This project focused on its application to dynamic models and relation to Kalman fileter. We provide a new approach to obtain the Kalman gain. A by-product of this project is to take the present Woodroofe-Stein ‘s identity one step further to obtain the multivariate Gram-Charlier series.

英文關鍵詞： Bayesian inference, dynamic systems, Kalman filter, Woodroofe-Stein ‘s identity

Final report for project: Bayesian inference for some dynamic systems

1 Introduction

The present project explored Woodroofe-Stein's identity and its applications to dynamic systems. First, it is shown that the current version of Woodroofe-Stein's identity can be taken one-step further to a series, from which a Gram-Charlier type expansion for multivariate densities can be obtained. This result has appeared [3]. Secondly, for the applications to dynamic systems, it is shown that this identity can give a novel derivation to the Kalman gain. The details are given in next section.

2 Kalman filter revisited

The Kalman filter [2] is a recursive method that estimate the latent state of a linear dynamic system. Consider the following linear dynamic model:

$$X_t = AX_{t-1} + W_{t-1} \tag{1}$$

$$Y_t = HX_t + V_t \tag{2}$$

where X_t is unobserved state variable and Y_t the measurement variable, both at time t. Suppose that X_t and W_t are n-dimensional vectors, Y_t and V_t are m-dimensional vectors, A is an $n \times n$ matrix, H is an $m \times n$ matrix, $W_t \sim N(0, Q)$, $V_t \sim N(0, R)$, Q is $n \times n$, R is $m \times m$, and both W_t and V_t are independent of X_t, A, H, Q, R are known. Let $D_t = \{Y_1, ..., Y_t\}$, the collection of data up to time t. The state of the system can be represented by the conditional mean and conditional covariance. Before observing the measurement Y_t, the *a priori* estimate of X_t and the error covariance matrix are defined as

$$\begin{aligned} \hat{X}_t^- &= E(X_t | D_{t-1}) \\ P_t^- &= E[(X_t - \hat{X}_t^-)(X_t - \hat{X}_t^-)^T | D_{t-1}]. \end{aligned} \tag{3}$$

Given knowledge of Y_t, the *a posteriori* state estimate and the *a posteriori* estimate error covariance are

$$\hat{X}_t = E(X_t | D_t) \tag{4}$$

$$P_t = E[(X_t - \hat{X}_t)(X_t - \hat{X}_t)^T | D_t), \tag{5}$$

Prediction step
$\hat{X}_t^- = A\hat{X}_{t-1}$
$P_t^- = AP_{t-1}A^T + Q$

Update step
$K_t = P_t^- H^T \left(HP_t^- H^T + R\right)^{-1}$
$\hat{X}_t = \hat{X}_t^- + K_t(Y_t - H\hat{X}_t^-)$
$P_t = (I_n - K_t H)P_t^-$

Table 1: Kalman filter

The Kalman filter propagates the state variable from time $t-1$ to time t. The filter consists of two steps: the *prediction* step and the *update* step. The prediction step infers from $(\hat{X}_{t-1}, P_{t-1})$ to $(\hat{X}_t^-, P_t^-)$, and the update step from $(\hat{X}_t^-, P_t^-)$ to $(\hat{X}_t, P_t)$. The prediction and update equations are given in Table 1. The $n \times m$ matrix K_t in the update step is called the Kalman gain. From the state equation (1) and the prior knowledge $(\hat{X}_{t-1}, P_{t-1})$ on X_{t-1}, the prediction equations can be easily obtained. The derivation of update equations are more complicated. It can be derived by first applying Bayes rule to the posterior density of X_t given D_t, expanding the numerator and denominator in the expression, rearranging all the terms, and employing the Matrix Inversion Lemma. Another approach starts by writing the *a posteriori* state estimate $\hat{X}_t$ as a linear combination of the *a priori* estimate $\hat{X}_t^-$ and the difference between the actual measurement Y_t and its prediction $H\hat{X}_t^-$,

$$\hat{X}_t = \hat{X}_t^- + K_t(Y_t - H\hat{X}_t^-), \tag{6}$$

and then determining the $n \times m$ matrix K_t by minimizing the mean-squared error

$$E((X_t - \hat{X}_t)^T(X_t - \hat{X}_t)|D_t).$$

The minimization involves substituting (6) into (5), differentiating the trace of P_t with respect to K_t, and setting the derivative equal to zero to solve K_t. The resulting K_t is

$$K_t = P_t^- H^T \left(HP_t^- H^T + R\right)^{-1}. \tag{7}$$

With the Kalman gain K_t, the posterior covariance matrix can be derived rather straightforwardly. To begin, write

$$\begin{aligned} P_t &= E[(X_t - \hat{X}_t)(X_t - \hat{X}_t)^T|D_t] \\ &= E\{[(X_t - \hat{X}_t^-) - K_t(HX_t + V_t - H\hat{X}_t^-)][(X_t - \hat{X}_t^-) - K_t(HX_t + V_t - H\hat{X}_t^-)]^T|D_t\} \\ &= P_t^- - P_t^- H^T K_t^T - K_t H P_t^- + K_t(HP_t^- H^T + R)K_t^T, \end{aligned} \tag{8}$$

where the last line follows from (3) and the independence of X_t and V_t. Then, substituting (7) into (8) gives the update of the error covariance estimate

$$P_t = (I - K_t H)P_t^-.$$

For details of this approach, see Brown and Hwang [1].

Now we show how to derive the update step by Stein's equation. First, write the posterior density of X_t given D_t as

$$\begin{aligned} p(x_t|D_t) &= p(x_t|y_t, D_{t-1}) \propto p(x_t|D_{t-1})p(y_t|x_t, D_{t-1}) \\ &= \text{prior } \times \text{likelihood} \\ &\propto \phi(x_t; \hat{x}_t^-, P_t^-)p(y_t|x_t). \end{aligned}$$

Next, let Σ_t satisfy $(P_t^-)^{-1} = \Sigma_t^T \Sigma_t$ and define

$$Z_t = \Sigma_t(X_t - \hat{X}_t^-). \tag{9}$$

So, $Z_t \sim N(0, I_n)$, where I_n is the $n \times n$ identity matrix. The posterior density of Z_t given D_t is

$$p(z_t|D_t) \propto f(z_t)\phi_n(z_t) \tag{10}$$

where

$$f(z_t) = e^{-\frac{1}{2}(y_t - H\Sigma_t^{-1}z_t - H\hat{x}_t^-)^T R^{-1}(y_t - H\Sigma_t^{-1}z_t - H\hat{x}_t^-)}. \tag{11}$$

The density (10) is of the form for Stein's equation. Therefore,

$$\begin{aligned} E(Z_t|D_t) &= E\left[\frac{\nabla_{z_t} f(Z_t)}{f(Z_t)}|D_t\right] \\ &= E\left((\Sigma_t^{-1})^T H^T R^{-1}(Y_t - H\Sigma_t^{-1}Z_t - H\hat{X}_t^-)|D_t\right); \end{aligned}$$

and collecting terms involving $E(Z_t|D_t)$ yields

$$\left(1 + (\Sigma_t^{-1})^T H^T R^{-1} H \Sigma_t^{-1}\right) E(Z_t|D_t) = (\Sigma_t^{-1})^T H^T R^{-1}(Y_t - H\hat{X}_t^-). \tag{12}$$

Now, from (4), (9), (12), and the property $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$, we obtain

$$\begin{aligned} \hat{X}_t &= E(X_t|D_t) = \hat{X}_t^- + \Sigma_t^{-1}E(Z_t|D_t) \\ &= \hat{X}_t^- + \Sigma_t^{-1}\left(1 + (\Sigma_t^{-1})^T H^T R^{-1} H \Sigma_t^{-1}\right)^{-1} (\Sigma_t^{-1})^T H^T R^{-1}(Y_t - H\hat{X}_t^-) \\ &= \hat{X}_t^- + \left((P_t^-)^{-1} + H^T R^{-1} H\right)^{-1} H^T R^{-1}(Y_t - H\hat{X}_t^-), \end{aligned}$$

which is of the form as (6). Then, the desired expression of the *a posteriori* state estimate can be derived by an application of the Matrix Inversion Lemma,

$$\left((P_t^-)^{-1} + H^T R^{-1} H\right)^{-1} H^T R^{-1} = P_t^- H^T \left(H P_t^- H^T + R\right)^{-1} \equiv K_t.$$

Once K_t is available, the posterior covariance matrix can be derived as shown in the previous paragraph.

3 Discussions

The result in the previous section is a joint work with Dr. Coad. It gives a new derivation for the Kalman gain. It will be combined with further findings and submitted later.

References

[1] R. G. Brown and P. Y. C. Hwang. *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley & Sons, Inc., New York, 3rd edition, 1997.

[2] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960.

[3] R. C. Weng. Expansions for multivariate densities. *Journal of Statistical Planning and Inference*, 167:174–181, 2015.

Report on attending The Joint Statistical Meeting, August 8 - 13, 2015, Seattle, Washington.

The Joint Statistical Meetings (JSM) is the largest statistical meeting held in North America. The JSM 2015 was held August 8-13, at the Washington State Convention Center. It was jointly held by organizations including American Statistical Association, Institute of Mathematical Statistics, International Chinese Statistical Association, International Society for Bayesian Analysis, Royal Statistical Society, etc. The conference puts together short courses, keynote lectures, scientific sessions, poster session, expositions, social events etc, and provides opportunities for participants to engage and network, and get inspirations to develop new ideas. This year it attracted over 6,000 participants.

This year I was invited by Professor X. Wang in Department of Statistics at University of Connecticut to the topic-contributed session sponsored by Bayesian Statistical Science. My talk title is "Real-time Bayesian inference for latent ability models." It is about Bayesian online inference for models such as paired-comparison models, item response theory models. I got a good chance to present my work and communicate with many people about my research. I also attended two professional short courses. One of them is "Applied text mining", which is a hands-on workshop with R code and packages for the practical application of text mining to real-world applications, including data from survey comments, websites, etc. The other one is "Software Engineering for Statisticians", which provides basics of computer architecture, revision control tools, code readability, etc. These materials are taught in a computer science curriculum, but seldom part of a statistics degree; however, they have become increasingly important tools for statisticians.

During these days, I met many old and new friends from industries and academics. Having chats with them inspired me and encouraged me to keep on moving. I have brought course materials from the two workshops. It was really a fruitful trip.

科技部補助計畫衍生研發成果推廣資料表

日期:2015/09/08

科技部補助計畫	計畫名稱: 動態系統的貝氏分析
	計畫主持人: 翁久幸
	計畫編號: 102-2118-M-004-003-MY2　　學門領域: 數理統計與機率
無研發成果推廣資料	

102年度專題研究計畫研究成果彙整表

計畫主持人：翁久幸　　計畫編號：102-2118-M-004-003-MY2

計畫名稱：動態系統的貝氏分析

成果項目			量化：實際已達成數（被接受或已發表）	量化：預期總達成數(含實際已達成數)	量化：本計畫實際貢獻百分比	單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
國內	論文著作	期刊論文	1	1	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	1	1	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	3	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	1	1	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	2	2	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

科技部補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

■達成目標

□未達成目標（請說明，以 100 字為限）

　　□實驗失敗

　　□因故實驗中斷

　　□其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文：□已發表 ■未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

動態系統與線上學習方法適用於資料量大且新資料不斷地增加的情況，當資料被使用之後,基於儲存等考量,可能不做保留,於當今資料爆炸的年代，這類方法更形重要.

本研究探討 Woodroofe-Stein 等式在這個領域之應用，已用此等式重新導出如卡爾曼濾波中的 Kalman gain； 並訓練學生對動態羅吉斯模型之認識與應用.