Pittsburgh Brain Activity Interpretation Competition 2006

Methods Description

Optimal Estimation of Brain Activation by Application of Kalman Filtering

Project Abstract

The feature rating time series from Movie1 and Movie2 was modeled as a vector AR(1) process, with state transition matrix **F**, and process noise covariance matrix **Q**. The Movie1 and Movie2 datasets were then used to estimate the measurement matrix **H** and the measurement covariance matrix **R**. This allowed implementation of a discrete time Kalman filter for estimation of the state vector (i.e., feature ratings) from the Movie3 data.

Introduction

The basic approach used was to treat the group of feature ratings as an unknown state vector, which could then be estimated using a discrete time Kalman filter. In order to implement the Kalman filter, it was necessary to estimate many parameters from the Movie1 and Movie2 data. The feature rating time series for Movie1 and Movie2 allowed the modeling of the feature rating as an AR(1) process, with state transition matrix **F**, and process noise covariance matrix **Q**. The Movie1 and Movie2 datasets were used to estimate the measurement matrix **H** and the measurement covariance matrix **R**. This allowed implementation of a discrete time Kalman filter for estimation of the state vector (i.e., feature ratings) from the Movie3 data.

Method

Pre-processing

The following pre-processing steps were applied to all datasets (i.e., Movie1, Movie2, and Movie3 fMRI data).

(1) The raw fMRI data was converted to AFNI format (Ref. Cox), using program to3d. This program assembles the slice data into 3D+time datasets.

(2) The fMRI data was corrected for subject motion, using AFNI program 3dvolreg, which performs a 6 dof rigid-body motion estimation and volume registration. In addition, the 6 dof estimated motion parameters were used as “nuisance” regressors in the next step.

(3) The next step is to remove any trends in the data (arising from scanner drift, etc.), and the removal of any remaining motion artifacts, on a voxel-by-voxel basis. This was done using the AFNI program 3dDeconvolve.

Parameter estimation

In order to determine which voxels were associated with the individual feature ratings, the above detrended data (for the Movie1 and Movie2 datasets) was used in a multiple linear regression analysis. The 30 volume and convolved feature rating time series were used as independent regressors in the multiple regression, performed using program 3dDeconvolve. The output of the program included the 30 regression parameters, measurement variance, and the F-statistic for the significance of the fit. Note that the regression parameters and measurement variance were used to initialize the **H** and **R** matrices, described below.

Thresholding based on full model F-statistic

In addition, the multiple regression full model F-statistic was used as a threshold to select voxels that were significantly correlated with the feature ratings. These "active" voxels were then used to extract data from the Movie3 dataset at the corresponding voxel locations.

Kalman filtering

A Matlab program was written to provide Kalman filtering of the above thresholded data extracted from the Movie3 dataset.

State model

The state vector $\mathbf{x}(k)$ represents the set of (p = 30) feature ratings at any point in time k.

$$\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ \vdots \\ x_{p-1}(k) \\ x_p(k) \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & F_{13} & \cdots & F_{1,p-1} & F_{1,p} \\ F_{21} & F_{22} & F_{23} & \cdots & F_{2,p-1} & F_{2,p} \\ F_{31} & F_{32} & F_{33} & \cdots & F_{3,p-1} & F_{3,p} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ F_{p-1,1} & F_{p-1,2} & F_{p-1,3} & \cdots & F_{p-1,p-1} & F_{p-1,p} \\ F_{p,1} & F_{p,2} & F_{p,3} & \cdots & F_{p,p-1} & F_{p,p} \end{bmatrix} \begin{bmatrix} x_1(k-1) \\ x_2(k-1) \\ x_3(k-1) \\ \vdots \\ x_{p-1}(k-1) \\ x_p(k-1) \end{bmatrix} + \begin{bmatrix} w_1(k) \\ w_2(k) \\ w_3(k) \\ \vdots \\ w_{p-1}(k) \\ w_p(k) \end{bmatrix}$$

or, in matrix notation:

$$\mathbf{x}(k) = \mathbf{F}\mathbf{x}(k-1) + w(k)$$

where $\mathbf{w}$(k) is iid N(0,$\mathbf{Q}$). Both the state transition matrix **F** and the process noise covariance matrix **Q** were estimated from the feature rating data provided for Movie1 and Movie2, by modeling the feature ratings as a vector AR(1) process.

Measurement model

The measurement model relates the fMRI measurement vector $\mathbf{z}$(k) (of length N = number of "active" voxels) to the state vector $\mathbf{x}$(k), thus:

$$
\begin{bmatrix} z_1(k) \\ z_2(k) \\ z_3(k) \\ \vdots \\ z_{N-1}(k) \\ z_N(k) \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} & \cdots & H_{1,p-1} & H_{1,p} \\ H_{21} & H_{22} & H_{23} & \cdots & H_{2,p-1} & H_{2,p} \\ H_{31} & H_{32} & H_{33} & \cdots & H_{3,p-1} & H_{3,p} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ H_{N-1,1} & H_{N-1,2} & H_{N-1,3} & \cdots & H_{N-1,p-1} & H_{N-1,p} \\ H_{N,1} & H_{N,2} & H_{N,3} & \cdots & H_{N,p-1} & H_{N,p} \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \\ \vdots \\ x_{p-1}(k) \\ x_p(k) \end{bmatrix} + \begin{bmatrix} v_1(k) \\ v_2(k) \\ v_3(k) \\ \vdots \\ v_{N-1}(k) \\ v_N(k) \end{bmatrix}
$$

or, in matrix notation:

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{v}(k)$$

where **v**(k) is iid N(0,**R**). Again, the **H** measurement matrix and the **R** measurement covariance matrix were estimated using multiple linear regression of the Movie1 and Movie2 data.

State and covariance prediction

This "state-space" formulation of the model allows implementation of the Kalman filter for optimal estimation of the state vector. Below, we summarize the discrete time Kalman filter equations. The predicted state vector $\hat{\mathbf{x}}(k\,|\,k-1)$ and the predicted error covariance matrix $\mathbf{P}(k\,|\,k-1)$ are given by:

$$
\begin{aligned}
\hat{\mathbf{x}}(k\,|\,k-1) &= \mathbf{F}\,\hat{\mathbf{x}}(k-1\,|\,k-1) \\
\mathbf{P}(k\,|\,k-1) &= \mathbf{F}\,\mathbf{P}(k-1\,|\,k-1)\,\mathbf{F}^T + \mathbf{Q}(k)
\end{aligned}
$$

From the above equations, we see that the previous state $\hat{\mathbf{x}}(k-1\,|\,k-1)$ and the previous state covariance matrix $\mathbf{P}(k-1\,|\,k-1)$ are predicted ahead by one time step, using the state transition matrix **F**.

State and covariance update

With each vector measurement **z**(k) (i.e., the fMRI data for the set of N voxels at time point k), the Kalman gain matrix $\mathbf{K}(k)$, updated state vector $\hat{\mathbf{x}}(k\,|\,k)$, and updated error covariance matrix $\mathbf{P}(k\,|\,k)$ are calculated by:

$$
\begin{aligned}
\mathbf{K}(k) &= \mathbf{P}(k\,|\,k-1)\,\mathbf{H}^T[\mathbf{H}\,\mathbf{P}(k\,|\,k-1)\,\mathbf{H}^T + \mathbf{R}(k)]^{-1} \\
\hat{\mathbf{x}}(k\,|\,k) &= \hat{\mathbf{x}}(k\,|\,k-1) + \mathbf{K}(k)[\mathbf{z}(k) - \mathbf{H}\,\hat{\mathbf{x}}(k\,|\,k-1)] \\
\mathbf{P}(k\,|\,k) &= \mathbf{P}(k\,|\,k-1) - \mathbf{K}(k)\,\mathbf{H}\,\mathbf{P}(k\,|\,k-1)
\end{aligned}
$$

The Kalman gain matrix **K**(k) is then calculated based on the relative sizes of the state covariance matrix $\mathbf{P}(k\,|\,k-1)$ and the measurement covariance matrix **R**. If **R** is large compared with $\mathbf{P}(k\,|\,k-1)$, i.e., if the uncertainty in the measurement is large compared with the uncertainty in the state estimate, then the Kalman gain **K**(k) will be small. However, if **R** is small compared with $\mathbf{P}(k\,|\,k-1)$, i.e., if the uncertainty in the measurement is small compared with the present uncertainty in the state estimate, then the Kalman gain **K**(k) will be large. In turn, the Kalman gain **K**(k) determines how much influence the next measurement **z**(k) has on the state estimate. If the Kalman gain is large, then the state estimate is greatly affected by the measurement. However, if the Kalman gain is small, then the measurement has little influence on the state estimate. The updated state vector and state covariance are then used as inputs to the next cycle of calculations.

References for these discrete-time Kalman filter equations include: Gelb; Bozic; and Srinath.

Results and Discussion

The feature rating time series for Movie1 and Movie2 allowed the modeling of the feature rating as an AR(1) process, with state transition matrix **F**, and process noise covariance matrix **Q**. The Movie1 and Movie2 datasets were used to estimate the measurement matrix **H** and the measurement covariance matrix **R**. This allowed implementation of a discrete time Kalman filter for estimation of the state vector (i.e., feature ratings) from the Movie3 data.

References

Cox, R. W. 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**:162-184.

Bozic, S.M. 1979. *Digital and Kalman Filtering*, Halsted Press.

Gelb A. (ed.) 19974. *Applied Optimal Estimation*, M.I.T. Press, Massachusets.

Srinath M., and Rajasekaran P. 1979. An Introduction to Statistical Signal Processing with Applications. Wiley-Interscience, New York.