Bayesian Change Point Detection for Bivariate Normal Data

1 Multiple Change Points Detection under the Bivariate Normal Model

This vignette demonstrates how to use the *bivnormcp* package to detect multiple change points in a sequence of bivariate normal observations. The methodology is Bayesian and assumes the bivariate data follows:

$$x_i = (x_{1i}, x_{2i})^\top \sim \mathcal{N}_2(\mu_i, \Sigma), \quad i = 1, \dots, n$$

,

where $\mu_i = (\mu_{1i} \;\; \mu_{2i})^\top$ and $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$

The probability density function of the vector $\mathbf{x}_i$ is given by -

$$\begin{aligned} & f(x_{1i}, x_{2i}|\mu_{1i}, \mu_{2i}, \rho) \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} exp\left[-\frac{1}{2(1-\rho^2)}\{(x_{1i}-\mu_{1i})^2 + (x_{2i}-\mu_{2i})^2 - 2\rho(x_{1i}-\mu_{1i})(x_{2i}-\mu_{2i})\}\right] \end{aligned}$$

Theorem 1.1 (Posterior Probability of Change Point) Let x_{it} denote the i^{th} component of the random vector $\mathbf{x}_t$ corresponding to the time point t, where $i = 1, 2$ and $t = 1, 2, \dots, T$. Assume that $\mathbf{x}_t = (x_{1t} \;\; x_{2t})^\top \sim N_2(\mu, \Sigma)$, where

$$\mu = (\mu_{1t} \;\; \mu_{2t})^\top \quad \text{and} \quad \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Let the joint prior distribution on the parameters, μ and Σ, be the Normal inverse Wishart (NIW) distribution with pdf given by

$$f(\mu, \Sigma|\mu_0, \lambda, \psi, \nu) = \frac{\lambda|\psi|^{\frac{\nu}{2}}|\Sigma|^{-\frac{\nu}{2}-2}}{(2\pi)2^\nu\Gamma_2(\frac{\nu}{2})} exp\left[-\frac{1}{2}tr(\psi\Sigma^{-1}) - \frac{\lambda}{2}(\mu-\mu_0)^T\Sigma^{-1}(\mu-\mu_0)\right],$$

where $\mu_0 = (\mu_{01} \;\; \mu_{02})^\top, \lambda, \psi, \nu$ are all hyper parameters. Then, the posterior probability of the change point being located at jth position at time-point (t+1) given data observed up to $t+1$, $\mathbf{x}_{1:t+1}$, is:

$$P(Q_{t+1} = j|\mathbf{x}_{1:t+1}) \propto \begin{cases} \omega^j_{t+1}\theta P(Q_t = j|\mathbf{x}_{1:t}) & \text{if } j < t \\ \omega^j_{t+1}(1-\theta)\sum_{i=0}^{t-1} P(Q_t = i|\mathbf{x}_{1:t}) & \text{if j = t} \end{cases}$$

where the weights are -

$$w^j_{t+1} \propto \begin{cases} \frac{t-j+\lambda}{t-j+\lambda+1} \frac{\int_{-1}^{1}(1-\rho^2)^{-\frac{\nu}{2}-2-\frac{t-j}{2}} exp\left[-\frac{1}{2(1-\rho^2)}(\eta_{11}-\eta_{12}-\eta_{13})\right] d\rho}{\int_{-1}^{1}(1-\rho^2)^{-\frac{\nu}{2}-2-\frac{t-j-1}{2}} exp\left[-\frac{1}{2(1-\rho^2)}(\eta_{21}-\eta_{22}-\eta_{23})\right] d\rho} & \text{if j < t} \\ \int_{-1}^{1} \frac{\lambda}{(\lambda+1)}(1-\rho^2)^{-\frac{\nu}{2}-2} exp\left[-\frac{1}{2(1-\rho^2)}(\eta_{31}-\eta_{32}-\eta_{33})\right] d\rho & \text{if j = t} \end{cases}$$

and,

$$\begin{aligned} \eta_{11} &= 2 + \sum_{i=j+1}^{t+1} x_{1i}^2 + \sum_{i=j+1}^{t+1} x_{2i}^2 - 2\rho \sum_{i=j+1}^{t+1} x_{1i}x_{2i} + \lambda\mu_{01}^2 - 2\lambda\mu_{01}\mu_{02}\rho + \lambda\mu_{02}^2, \\ \eta_{12} &= \frac{\left(\lambda\mu_{01} - \lambda\mu_{02}\rho + \sum_{i=j+1}^{t+1} x_{1i} - \rho\sum_{i=j+1}^{t+1} x_{2i}\right)^2}{t-j+\lambda+1}, \\ \eta_{13} &= \frac{(1-\rho^2)\left(\lambda\mu_{02} + \sum_{i=j+1}^{t+1} x_{2i}\right)^2}{t-j+\lambda+1}, \\ \eta_{21} &= 2 + \sum_{i=j+1}^{t} x_{1i}^2 + \sum_{i=j+1}^{t} x_{2i}^2 - 2\rho \sum_{i=j+1}^{t} x_{1i}x_{2i} + \lambda\mu_{01}^2 - 2\lambda\mu_{01}\mu_{02}\rho + \lambda\mu_{02}^2, \\ \eta_{22} &= \frac{\left(\lambda\mu_{01} - \lambda\mu_{02}\rho + \sum_{i=j+1}^{t} x_{1i} - \rho\sum_{i=j+1}^{t} x_{2i}\right)^2}{t-j+\lambda}, \\ \eta_{23} &= \frac{(1-\rho^2)\left(\lambda\mu_{02} + \sum_{i=j+1}^{t} x_{2i}\right)^2}{t-j+\lambda}, \\ \eta_{31} &= 2 + x_{1,t+1}^2 + x_{2,t+1}^2 - 2\rho x_{1,t+1}x_{2,t+1} + \lambda\mu_{01}^2 - 2\lambda\mu_{01}\mu_{02}\rho + \lambda\mu_{02}^2, \\ \eta_{32} &= \frac{(\lambda\mu_{01} - \lambda\mu_{02}\rho + x_{1,t+1} - \rho x_{2,t+1})^2}{\lambda+1}, \\ \eta_{33} &= \frac{(1-\rho^2)(\lambda\mu_{02} + x_{2,t+1})^2}{\lambda+1}. \end{aligned}$$

2 Usage

2.1 Introduction

This vignette demonstrates how to use the *bivnormcp* package to detect change points in bivariate normal data and visualize them.

2.2 Simulated Example

We'll create a synthetic dataset with two change points:

```
set.seed(123)

test_data <- data.frame(
  var1 = c(rnorm(50, 0, 0.5), rnorm(25, 5, 0.5), rnorm(25, 15, 0.5)),
  var2 = c(rnorm(50, 0, 0.5), rnorm(25, 5, 0.5), rnorm(25, 15, 0.5))
)
```

2.3 Run Change Point Detection

```
result <- bivar_norm_cp(test_data$var1, test_data$var2,
                        th_cp = 0.9,
                        save_output = FALSE)  # we won't save during vignette build
result
#> $x1
#>   [1] -0.28023782 -0.11508874  0.77935416  0.03525420  0.06464387  0.85753249
#>   [7]  0.23045810 -0.63253062 -0.34342643 -0.22283099  0.61204090  0.17990691
#>  [13]  0.20038573  0.05534136 -0.27792057  0.89345657  0.24892524 -0.98330858
#>  [19]  0.35067795 -0.23639570 -0.53391185 -0.10898746 -0.51300222 -0.36444561
#>  [25] -0.31251963 -0.84334666  0.41889352  0.07668656 -0.56906847  0.62690746
#>  [31]  0.21323211 -0.14753574  0.44756283  0.43906674  0.41079054  0.34432013
#>  [37]  0.27695883 -0.03095586 -0.15298133 -0.19023550 -0.34735349 -0.10395864
#>  [43] -0.63269818  1.08447798  0.60398100 -0.56155429 -0.20144242 -0.23332768
#>  [49]  0.38998256 -0.04168453  5.12665926  4.98572662  4.97856477  5.68430114
#>  [55]  4.88711451  5.75823530  4.22562360  5.29230687  5.06192712  5.10797078
#>  [61]  5.18981974  4.74883827  4.83339631  4.49071231  4.46410439  5.15176432
#>  [67]  5.22410489  5.02650211  5.46113373  6.02504234  4.75448442  3.84541556
#>  [73]  5.50286926  4.64539962  4.65599569 15.51278568 14.85761350 14.38964114
#>  [79] 15.09065174 14.93055432 15.00288209 15.19264020 14.81466998 15.32218827
#>  [85] 14.88975672 15.16589098 15.54841951 15.21759075 14.83703421 15.57440381
#>  [91] 15.49675193 15.27419848 15.11936587 14.68604696 15.68032622 14.69987021
#>  [97] 16.09366650 15.76630531 14.88214982 14.48678955
#> 
#> $x2
#>   [1] -0.35520328  0.12844185 -0.12334594 -0.17377130 -0.47580928 -0.02251386
#>   [7] -0.39245223 -0.83397097 -0.19011326  0.45949830 -0.28767348  0.30398216
#>  [13] -0.80894135 -0.02778098  0.25970360  0.15057668  0.05283810 -0.32035300
#>  [19] -0.42485217 -0.51206440  0.05882330 -0.47373731 -0.24527872 -0.12804610
#>  [25]  0.92193100 -0.32597495  0.11769329  0.03898042 -0.48092832 -0.03565404
#>  [31]  0.72227543  0.22575203  0.02061646 -0.21124842 -1.02662361  0.56566861
#>  [37] -0.73032004  0.36997376  0.95455178 -0.72194658  0.35089217 -0.13109874
#>  [43] -0.78607208 -0.75733383 -0.80076809 -0.26545326 -0.73087779  0.34395839
#>  [49]  1.05005447 -0.64351524  5.39386942  5.38452112  5.16610129  4.49581170
#>  [55]  4.94027370  4.85980233  5.28149477  4.81378062  5.48848669  4.81270957
#>  [61]  5.52635573  4.47541150  4.36992238  6.62051997  4.79157121  5.14911380
#>  [67]  5.31828484  4.75810969  5.25843102  5.18448226  4.89230975  5.03264652
#>  [73]  4.98296637  6.06422595  4.62933195 14.45200187 15.01889420 15.15524037
#>  [79] 15.21826174 14.77081733 14.46833693 15.63159259 14.82517481 14.56724357
#>  [85] 14.88186022 14.90141205 15.55496014 15.04236865 15.37702689 14.75035399
#>  [91] 15.10722265 14.83765704 15.04729176 14.55231832 14.34459923 15.99860669
#>  [97] 15.30035441 14.37436432 14.69441704 14.40725996
#> 
#> $prob_max
#>            [,1] [,2]
#>   [1,] 0.9979290    0
#>   [2,] 0.9945776    0
#>   [3,] 0.9956605    0
#>   [4,] 0.9955205    0
#>   [5,] 0.9935281    0
#>   [6,] 0.9941091    0
#>   [7,] 0.9942935    0
#>   [8,] 0.9944815    0
#>   [9,] 0.9897565    0
#>  [10,] 0.9917490    0
#>  [11,] 0.9920558    0
#>  [12,] 0.9921507    0
#>  [13,] 0.9932116    0
#>  [14,] 0.9900119    0
#>  [15,] 0.9919504    0
#>  [16,] 0.9930271    0
#>  [17,] 0.9910173    0
#>  [18,] 0.9932876    0
#>  [19,] 0.9941956    0
#>  [20,] 0.9913394    0
#>  [21,] 0.9928226    0
#>  [22,] 0.9918897    0
#>  [23,] 0.9910539    0
#>  [24,] 0.9743651    0
#>  [25,] 0.9777765    0
#>  [26,] 0.9878315    0
#>  [27,] 0.9896699    0
#>  [28,] 0.9904922    0
#>  [29,] 0.9921667    0
#>  [30,] 0.9908980    0
#>  [31,] 0.9905281    0
#>  [32,] 0.9917860    0
#>  [33,] 0.9918057    0
#>  [34,] 0.9789903    0
#>  [35,] 0.9895712    0
#>  [36,] 0.9871207    0
#>  [37,] 0.9886225    0
#>  [38,] 0.9737070    0
#>  [39,] 0.9877613    0
#>  [40,] 0.9842745    0
#>  [41,] 0.9865671    0
#>  [42,] 0.9912637    0
#>  [43,] 0.9654154    0
#>  [44,] 0.9547677    0
#>  [45,] 0.9711169    0
#>  [46,] 0.9723217    0
#>  [47,] 0.9755005    0
#>  [48,] 0.9779766    0
#>  [49,] 0.9846031    0
#>  [50,] 0.9997747   50
#>  [51,] 0.9988292   50
#>  [52,] 0.9983230   50
#>  [53,] 0.9954243   50
#>  [54,] 0.9964824   50
#>  [55,] 0.9961886   50
#>  [56,] 0.9956868   50
#>  [57,] 0.9968274   50
#>  [58,] 0.9969912   50
#>  [59,] 0.9970812   50
#>  [60,] 0.9967899   50
#>  [61,] 0.9936516   50
#>  [62,] 0.9903377   50
#>  [63,] 0.9905237   50
#>  [64,] 0.9928362   50
#>  [65,] 0.9947368   50
#>  [66,] 0.9949995   50
#>  [67,] 0.9953091   50
#>  [68,] 0.9946898   50
#>  [69,] 0.9908735   50
#>  [70,] 0.9925112   50
#>  [71,] 0.9884865   50
#>  [72,] 0.9932846   50
#>  [73,] 0.9944684   50
#>  [74,] 0.9936203   50
#>  [75,] 1.0000000   75
#>  [76,] 0.9993403   75
#>  [77,] 0.9982788   75
#>  [78,] 0.9983500   75
#>  [79,] 0.9985190   75
#>  [80,] 0.9987338   75
#>  [81,] 0.9985198   75
#>  [82,] 0.9984294   75
#>  [83,] 0.9986279   75
#>  [84,] 0.9986416   75
#>  [85,] 0.9986126   75
#>  [86,] 0.9984997   75
#>  [87,] 0.9984373   75
#>  [88,] 0.9982241   75
#>  [89,] 0.9984166   75
#>  [90,] 0.9984875   75
#>  [91,] 0.9984806   75
#>  [92,] 0.9984883   75
#>  [93,] 0.9984279   75
#>  [94,] 0.9973203   75
#>  [95,] 0.9954723   75
#>  [96,] 0.9976614   75
#>  [97,] 0.9960227   75
#>  [98,] 0.9975193   75
#>  [99,] 0.9980058   75
#> 
#> $series_length
#> [1] 100
```

2.4 Correct the Indices and Visualize Change Points

```
change_pts <- index_correction(result$prob_max[, 2])
plot_cp_3d(result$x1, result$x2, change_pts)
#> Warning: 'layout' objects don't have these attributes: 'NA'
#> Valid attributes include:
#> '_deprecated', 'activeshape', 'annotations', 'autosize', 'autotypenumbers', 'calendar', 'clickmode', 'coloraxi
s', 'colorscale', 'colorway', 'computed', 'datarevision', 'dragmode', 'editrevision', 'editType', 'font', 'geo',
'grid', 'height', 'hidesources', 'hoverdistance', 'hoverlabel', 'hovermode', 'images', 'legend', 'mapbox', 'margi
n', 'meta', 'metasrc', 'modebar', 'newshape', 'paper_bgcolor', 'plot_bgcolor', 'polar', 'scene', 'selectdirectio
n', 'selectionrevision', 'separators', 'shapes', 'showlegend', 'sliders', 'smith', 'spikedistance', 'template',
'ternary', 'title', 'transition', 'uirevision', 'uniformtext', 'updatemenus', 'width', 'xaxis', 'yaxis', 'barmod
e', 'bargap', 'mapType'
```