國立交通大學

統計學研究所

碩士論文

利用計分檢定統計量之指數加權移動平均控制圖來監控一般線性輪廓

An Exponentially Weighted Moving Average Control Chart Based on Score Test Statistics for Monitoring General Linear Profiles

研 究 生：馮千育

指導教授：陳志榮 博士

中 華 民 國 一百零二 年 四 月

利用計分檢定統計量之指數加權移動平均控制圖來監控一般線性輪廓

An Exponentially Weighted Moving Average Control Chart Based on Score Test Statistics for Monitoring General Linear Profiles

研 究 生：馮千育　　Student：Chien-Yu Feng

指導教授：陳志榮 博士　　Advisor：Dr. Chih-Rung Chen

國 立 交 通 大 學 理 學 院

統 計 學 研 究 所

碩 士 論 文

A Thesis

Submitted to Institute of Statistics

College of Science National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Statistics

April 2013

Hsinchu, Taiwan, Republic of China

中華民國一百零二年四月

利用計分檢定統計量之指數加權移動平均控制圖來監控一般線性輪廓

研究生：馮千育　　　　　　　　　　　　　　　　　　　指導教授:陳志榮 博士

國立交通大學統計學研究所碩士班

摘　　要

統計製程控制是常常用在監控工業實務的一個方法,其中控制圖更是一個用重要的監控工具。本文提出利用計分檢定統計量的指數加權移動平均來監控一般線性輪廓的資料。並且用模擬的方法來跟一些文獻的資料做比較，來探討它的表現與優劣。

關鍵字: 計分檢定統計量、指數加權移動平均、一般線性輪廓

An Exponentially Weighted Moving Average Control Chart Based on Score Test Statistics for Monitoring General Linear Profiles

Student：Chien-Yu Feng　　　　Advisor：Dr. Chih-Rung Chen

Institute of Statistics

National Chiao Tung University

Abstract

Statistical process control is often utilized to monitor an industrial process. Control charts are important monitoring tools used to determine whether a process is in a state of statistical control. An exponentially weighted moving average control chart based on score test statistics to monitor general linear profiles is proposed in this paper. The performance of the proposed monitoring scheme is compared with references through a simulation study.

KEY WORDS: Score test statistics, Exponentially weighted moving average control chart, General linear profiles

誌　　謝

我的研究所生涯，比別人多的快一年的時間，一切都是自己平時不認真的後果，現在終於能夠畢業了，實在非常感動。

首先我非常感謝我的指導教授-陳志榮老師，從二年級開始一直到我畢業，比其他同學多了快一年的時間，老師依然未間斷的辛苦教導，無論怎樣的問題老師都會詳細的跟我解說，並教導我嚴謹的研究態度，老師對我的照顧，無限感激。此外也感謝替我口試的口試委員們，感謝蔡明田教授特別遠從臺北下來新竹一趟，感謝許文郁教授感冒還能特別前來，感謝洪志真教授給我論文的建議。

最後要感謝交大統計所的所有老師與工作人員們，讓我茫然無助時能夠指引我學習研究的方向。特別還要感謝我的家人們，如不是你們一直以來的包容及支持，我也不可能有今天的成就。

馮千育　謹誌于

國立交通大學統計學研究所

中華民國一百零二年四月

Contents

List of Tables

1 Introduction

1.1 Motivation

Control charts are important monitoring tools used to determine whether a manufacturing or business process is in a state of statistical control. Roberts (1959) proposed an exponentially weighted moving average (EWMA) control chart which is sensitive to a small shift in the in-control process mean. The likelihood ratio (LR) test is a statistical test used to compare the fit of two models.

Monitoring schemes by an EWMA control chart based on the LR test statistics have been proposed in the literature, e.g., Zou *et al.* (2006) proposed an EWMA control chart based on LR test statistics for monitoring simple linear profiles with unknown in-control process parameters and Zou *et al.* (2007) proposed a multivariate EWMA (MEWMA) control chart based on LR test statistics for monitoring general linear profiles with known in-control process parameters. The LR and score test statistics are asymptotically equivalent under the null hypothesis. Sometimes, it is easier to evaluate the score test statistic than the LR test statistic. The in-control process parameters are usually unknown in practice. We would like to utilize the score test statistics to monitor general linear profiles by an EWMA monitoring scheme, and then to investigate the performance of the proposed methodology.

1.2 Literature Review

Statistical process control (SPC) is often utilized to monitor an industrial process. How to construct a control chart is an important issue in SPC. The control chart is used to determine whether a manufacturing or business process is in a state of statistical control.

The control chart can be used in both phases I and II. In phase I, some reference data are collected and analyzed to assess whether they are in control. Then the in-control process parameters and control limits are estimated from the in-control data identified from those reference data. In phase II, the process is monitored over time to see whether it is in control by using control limits. The average run length (ARL) is usually used to appraise the process performance.

Shewhart (1931) proposed the $\bar{X}$ control chart which has been used to monitor the process mean and has good performance for a large sample size or for detecting a large shift in the process mean.

Page (1954) proposed the cumulative sum (CUSUM) control chart whose performance is better than that of the Shewhart control chart in detecting a small sustained shift in the process mean.

The EWMA control chart was proposed by Roberts (1959) for detecting a small sustained shift in the process mean. Its performance for detecting a small sustained shift in the process mean is better than that of the Shewhart control chart. As the in-control process parameters are usually unknown in practice, Jones *et al.* (2001) investigated the performance of the EWMA control chart utilizing estimated in-control process parameters and derived its run-length (RL) distribution. Castagliola *et al.* (2006) reviewed the EWMA control chart for monitoring the process position and variability. Jensen *et al.* (2006) reviewed the effect of parameter estimation and proposed some recommendations for future research.

Sometimes, we are interested in the relationship between a response variable and one or more explanatory variables in the process. Kim *et al.* (2003) proposed a method based on three EWMA control charts, where these three charts were used for different process parameters in simple linear profiles assuming the in-control process parameters are known. Zou *et al.* (2006) proposed an LR-based control chart for a change-point model to monitor simple linear profiles assuming the in-control process parameters are unknown. Zou *et al.* (2007) proposed an MEWMA control chart for monitoring general linear profiles assuming the in-control process parameters are known. Zou *et al.* (2009) compared five control schemes for monitoring the process mean subject to drifts. Zou *et al.* (2010) proposed a single chart that integrated the EWMA procedure with the LR test statistics for monitoring both the process mean and variance. Huang (2012) proposed an EWMA control chart based on LR test statistics for monitoring general linear profiles.

Kim *et al.* (2003) proposed three EWMA control charts for monitoring simple linear profiles as follows: Suppose that data $\{(x_i, y_{ij}) : \ i = 1, 2, \ldots, n\}$ are available at time $j = 1, 2, \ldots, \tau$, where x_is are not all the same and an out-of-control signal occurs at time τ. The process is called in control at time j if

$$\boldsymbol{y}_{ij} = \beta_0 + \beta_1 x_i + \sigma\,\varepsilon_{ij}, \qquad i = 1, 2, \ldots, n, \tag{1.1}$$

where ε_{ij}s are independent standard normal random variables. Model (1.1) is equivalent to

$$\boldsymbol{y}_{ij} = \beta_0' + \beta_1' x_i' + \sigma\,\varepsilon_{ij}, \qquad i = 1, 2, \ldots, n, \tag{1.2}$$

where $\beta_0' = \beta_0 + \beta_1 \bar{x}$, $\beta_1' = \beta_1$, and $x_i' = x_i - \bar{x}$ with $\bar{x} = \Sigma_{i=1}^{n} x_i/n$. At time j, the least-squares estimator of β_0', β_1' , and σ^2 are

$$b_{0j} = \bar{\boldsymbol{y}}_j,$$
$$b_{1j} = \frac{\Sigma_{i=1}^{n}(x_i - \bar{x})\boldsymbol{y}_{ij}}{\Sigma_{i=1}^{n}(x_i - \bar{x})^2},$$

and

$$\mathrm{MSE}_j = \frac{1}{n-2}\sum_{i=1}^{n}(\boldsymbol{y}_{ij} - b_{0j} - b_{1j}x_i')^2,$$

where $\bar{\boldsymbol{y}}_j = \Sigma_{i=1}^{n} y_i/n$. Since b_{0j}, b_{1j}, and MSE_j are independent random variables, they proposed three EWMA control charts

$$\mathrm{EWMA_I}(j) = \lambda b_{0j} + (1-\lambda)\mathrm{EWMA_I}(j-1),$$
$$\mathrm{EWMA_S}(j) = \lambda b_{1j} + (1-\lambda)\mathrm{EWMA_S}(j-1),$$

and

$$\mathrm{EWMA_E}(j) = \max\{\lambda \ln(\mathrm{MSE}_j) + (1-\lambda)\mathrm{EWMA_E}(j-1), \ln(\sigma^2)\},$$

where $\mathrm{EWMA_I}(0) = \beta_0'$, $\mathrm{EWMA_S}(0) = \beta_1'$, $\mathrm{EWMA_E}(0) = \ln(\sigma^2)$, and λ is a smoothing parameter in (0, 1]. Those three EWMA control charts are proposed to monitor β_0', β_1', and σ, respectively, using the same in-control ARL for each control chart.

Zou *et al.* (2007) proposed an MEWMA control chart for monitoring general linear profiles as follows: Suppose that data $(X,\ \boldsymbol{y}_j)$ are available at time $j = 1, 2, \ldots, \tau$, where an out-of-control signal occurs at time τ. The process is called in control at time j if

$$\boldsymbol{y}_j = X\boldsymbol{\beta} + \sigma\,\varepsilon_j, \quad j = 1, 2, \ldots, \tau, \tag{1.3}$$

where $\boldsymbol{y}_j$ is an $n \times 1$ response vector and X is an $n \times p$ known model (or design) matrix of rank p $(< n_j)$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})^T$ is a known in-control $p \times 1$ process regression parameter vector, σ is a known in-control positive process scale parameter, and ε_js are independent standardized error vectors with $\varepsilon_j \sim N_n(0_{n\times 1}, I_n)$. Set

$$\mathbf{Z}_j(\boldsymbol{\beta}) = \frac{\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}}{\sigma}$$

and

$$\mathbf{Z}_j(\sigma) = \Phi^{-1}\left(F_{\chi^2_{n-p}}\left(\frac{(n-p)\hat{\sigma}_j^2}{\sigma^2}\right)\right),$$

where $\hat{\boldsymbol{\beta}}_j = (X^TX)^{-1}X^T\boldsymbol{y}_j$, $\hat{\sigma}_j^2 = (\boldsymbol{y}_j - X\hat{\boldsymbol{\beta}}_j)^T(\boldsymbol{y}_j - X\hat{\boldsymbol{\beta}}_j)/(n-p)$, $\Phi^{-1}(\cdot)$ is the inverse function of the standard normal cumulative distribution function (c.d.f.), and $F_{\chi^2_{n-p}}(\cdot)$ is the chi-squared c.d.f. with $n - p$ degrees of freedom. Set $\mathbf{Z}_j \equiv (\mathbf{Z}_j^T(\boldsymbol{\beta}), \mathbf{Z}_j(\sigma))^T$, a $(p+1) \times 1$ random vector. When the process is in control at time j, $\mathbf{Z}_j$ is multivariate normally distributed with mean vector $0_{(p+1)\times 1}$ and covariance matrix

$$\Sigma = \begin{pmatrix} (X^TX)^{-1} & 0_{p\times 1} \\ 0_{1\times p} & 1 \end{pmatrix}_{(p+1)\times(p+1)}.$$

Then the MEWMA sequence is defined as

$$\mathbf{W}_j = \lambda \mathbf{Z}_j + (1-\lambda)\mathbf{W}_{j-1}, \qquad j = 1, 2, \ldots,$$

where $\mathbf{W}_0 \equiv 0_{(p+1)\times 1}$ and λ is a smoothing parameter in $(0, 1]$. An out-of-control signal occurs at time j if

$$\mathbf{W}_j \Sigma^{-1} \mathbf{W}_j > L \frac{\lambda}{2-\lambda},$$

where L (> 0) is chosen to achieve a specified in-control ARL.

In Section 2, general linear profiles are described and then an EWMA control chart based on score test statistics is proposed for monitoring general linear profiles. In Section 3, a simulation study is presented to illustrate the proposed methodology. In Section 4, conclusions are given. In Section 5, some potential future works is suggested.

2 An EWMA control chart for monitoring general linear profiles

In this section, general linear profiles are described and then an EWMA control chart based on score test statistics is proposed for monitoring linear profiles.

2.1 Model

Suppose that data $\{(y_{ij}, x_{ij})\colon\ i = 1, 2, \ldots, n_j\}$ are available at time $j = 1, 2, \ldots, \tau$, where y_{ij} is the ith response variable at time j, x_{ij} is its corresponding explanatory variable(s), and an out-of-control signal occurs at time τ. Assume that

$$y_{ij} = \beta_{j0}\, u_0(x_{ij}) + \beta_{j1}\, u_1(x_{ij}) + \cdots + \beta_{j,p-1}\, u_{p-1}(x_{ij}) + \sigma_j\, \varepsilon_{ij}, \qquad i = 1, 2, \ldots, n_j, \quad (2.1)$$

where $\beta_{j0}, \beta_{j1}, \ldots, \beta_{j,p-1}$ are unknown real-valued process regression parameters at time j; $u_0(\cdot), u_1(\cdot), \ldots, u_{p-1}(\cdot)$ are known real-valued functions; σ_j is an unknown positive process scale parameter at time j; and ε_{ij}s are *i.i.d.* $N(0, 1)$ standardized errors.

Example 1: Model (2.1) has the form

$$y_{ij} = \beta_{j0} + \beta_{j1}\, x_{ij} + \cdots + \beta_{j,p-1}\, x_{ij}^{p-1} + \sigma_j\, \varepsilon_{ij}, \qquad i = 1, 2, \ldots, n_j,$$

for simple linear profiles if $p = 2$ or for polynomial profiles if $p \geq 3$.

Example 2: Model (2.1) has the form

$$y_{ij} = \beta_{j0} + \sum_{u=1}^{k} \beta_{ju} x_{iju} + \sum_{u=1}^{k} \beta_{juu} x_{iju}^2 + \sum_{1 \leq u < u' \leq k} \beta_{juu'} x_{iju} x_{iju'} + \sigma_j \varepsilon_{ij}, \qquad i = 1, 2, \ldots, n_j,$$

with $x_{ij} \equiv (x_{ij1}, \ldots, x_{ijk})^T$ for quadratic polynomial profiles if $k \geq 2$.

For simplicity of notation, model (2.1) is rewritten as

$$\boldsymbol{y}_j = X_j \boldsymbol{\beta}_j + \sigma_j \varepsilon_j, \tag{2.2}$$

where $\boldsymbol{y}_j$ ($\equiv (y_{j1}, y_{j2}, ..., y_{jn_j})^T$) is an $n_j \times 1$ response vector at time j, X_j is an known $n_j \times p$ model (or design) matrix of full rank p ($< n_j$) at time j, $\boldsymbol{\beta}_j$ ($\equiv (\beta_{j0}, \beta_{j1}, \ldots, \beta_{j,p-1})^T$) is a $p \times 1$ parameter vector of unknown real-valued process regression parameters at time j, σ_j is an unknown positive process scale parameter at time j, and ε_js are independent standardized error vectors with $\varepsilon_j \sim N_{n_j}(0_{n_j \times 1}, I_{n_j})$. Set $\boldsymbol{\theta}_j \equiv (\boldsymbol{\beta}_j^T, \sigma_j)^T$ ($\in \mathcal{R}^p \times (0, \infty) \equiv \Theta$), the process parameter vector at time j. Set $\boldsymbol{\theta} \equiv (\boldsymbol{\beta}^T, \sigma)^T$ ($\in \Theta$), the in-control process parameter vector.

2.2 Known in-control process parameters without constraint

In this subsection, assume that the in-control process parameter vector $\boldsymbol{\theta}$ is known. The process is called in control at time j if $\boldsymbol{\theta}_j = \boldsymbol{\theta}$ or out of control at time j if $\boldsymbol{\theta}_j \neq \boldsymbol{\theta}$.

For model (2.2), the joint probability density function (p.d.f.) of $\boldsymbol{y}_j$ at time j is

$$
\begin{aligned}
f(\boldsymbol{y}_j;\boldsymbol{\theta}_j) &= \frac{1}{(2\pi)^{n_j/2}\sigma_j^{n_j}} \exp\left\{-\frac{\|\boldsymbol{y}_j - X_j\boldsymbol{\beta}_j\|^2}{2\sigma_j^2}\right\} \\
&= \frac{1}{(2\pi)^{n_j/2}\sigma_j^{n_j}} \exp\left\{-\frac{(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j)^T X_j^T X_j(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j) + n_j\hat{\sigma}_j^2}{2\sigma_j^2}\right\},
\end{aligned}
$$

where $(\hat{\boldsymbol{\beta}}_j^T, \hat{\sigma}_j)^T$ $(\equiv \hat{\boldsymbol{\theta}}_j)$ is the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}_j$ such that $\hat{\boldsymbol{\beta}}_j$ is independent of $\hat{\sigma}_j$ with

$$
\hat{\boldsymbol{\beta}}_j \equiv (X_j^T X)_j^{-1} X_j^T \boldsymbol{y}_j \sim N_p(\boldsymbol{\beta}_j, \sigma_j^2 (X_j^T X_j)^{-1}) \tag{2.3}
$$

and

$$
\hat{\sigma}_j^2 \equiv \frac{\|\boldsymbol{y}_j - X_j\hat{\boldsymbol{\beta}}_j\|^2}{n_j} = \frac{\left\|\left[I_{n_j} - X_j(X_j^T X_j)^{-1}X_j^T\right]\boldsymbol{y}_j\right\|^2}{n_j} \sim \frac{\sigma_j^2}{n_j}\chi^2_{n_j-p}. \tag{2.4}
$$

Then the log-likelihood function for $\boldsymbol{\theta}_j$ at time j is

$$
\ell_j(\boldsymbol{\theta}_j) \equiv \log\left[f(\boldsymbol{y}_j;\boldsymbol{\theta}_j)\right] = -\frac{n_j}{2}\log(2\pi) - \frac{n_j}{2}\log(\sigma_j^2) - \frac{\|\boldsymbol{y}_j - X_j\boldsymbol{\beta}_j\|^2}{2\sigma_j^2}.
$$

The corresponding LR and score test statistics at time j are

$$
2\left[\ell_j(\hat{\boldsymbol{\theta}}_j) - \ell_j(\boldsymbol{\theta})\right]
$$

and

$$
\left.\frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j^T} \mathrm{Cov}_{\boldsymbol{\theta}_j}^{-1}\left(\frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j}\right)\frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j}\right|_{\boldsymbol{\theta}_j=\boldsymbol{\theta}} (\equiv W_j),
$$

respectively, where $\partial \ell_j(\boldsymbol{\theta}_j)/\partial \boldsymbol{\theta}_j$ is the score function for $\boldsymbol{\theta}_j$ at time j such that

$$\frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\beta}_j} = \frac{X_j^T(\boldsymbol{y}_j - X_j\boldsymbol{\beta}_j)}{\sigma_j^2} \sim N_p\left(0_{p\times 1}, \frac{X_j^T X_j}{\sigma_j^2}\right),$$
$$\frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \sigma_j} = \frac{\|\boldsymbol{y}_j - X_j\boldsymbol{\beta}_j\|^2}{\sigma_j^3} - \frac{n_j}{\sigma_j} \sim \frac{\chi^2_{n_j} - n_j}{\sigma_j},$$

and $\partial \ell_j(\boldsymbol{\theta}_j)/\partial \boldsymbol{\beta}_j$ is independent of $\partial \ell_j(\boldsymbol{\theta}_j)/\partial \sigma_j$.

When the process is in control at time j,

$$2\left[\ell_j(\hat{\boldsymbol{\theta}}_j) - \ell_j(\boldsymbol{\theta})\right] = W_j + O_p\left(\frac{1}{\sqrt{n_j}}\right) \tag{2.5}$$

as $n_j \to \infty$, where equation (2.5) is sketched in Appendix A.1.

$$\begin{aligned} W_j &= \frac{(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})^T X_j^T X_j (\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})}{\sigma^2} + \frac{1}{2n_j}\left[\frac{(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})^T X_j^T X_j (\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})}{\sigma^2} + \frac{n_j\hat{\sigma}_j^2}{\sigma^2} - n_j\right]^2 \\ &\equiv H_{j1} + \frac{1}{2n_j}(H_{j1} + H_{j2} - n_j)^2, \end{aligned} \tag{2.6}$$

where H_{j1} is independent of H_{j2} such that

$$H_{j1} \equiv \frac{(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})^T X_j^T X_j (\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})}{\sigma^2} \sim \frac{\sigma_j^2}{\sigma^2}\chi_p^2((\boldsymbol{\beta}_j - \boldsymbol{\beta})^T X_j^T X_j(\boldsymbol{\beta}_j - \boldsymbol{\beta})/\sigma_j^2) \tag{2.7}$$

and

$$H_{j2} \equiv \frac{n_j\hat{\sigma}_j^2}{\sigma^2} \sim \frac{\sigma_j^2}{\sigma^2}\chi^2_{n_j-p} \tag{2.8}$$

with $(\boldsymbol{\beta}_j - \boldsymbol{\beta}_0)^T X_j^T X_j(\boldsymbol{\beta}_j - \boldsymbol{\beta}_0)/\sigma_j^2$ being the noncentrality parameter for the noncentral

χ_p^2 distribution if $(\boldsymbol{\beta}_j - \boldsymbol{\beta}_0)^T X_j^T X_j (\boldsymbol{\beta}_j - \boldsymbol{\beta}_0)/\sigma_j^2 > 0$. Set

$$\tau_j^2 \equiv \frac{(\boldsymbol{\beta}_j - \boldsymbol{\beta})^T X_j^T X_j (\boldsymbol{\beta}_j - \boldsymbol{\beta})}{\sigma_j^2} \ (\geq 0).$$

Note that the distribution of W_j depends only on all of p, n_j, σ_j/σ, and τ_j^2. When the process is in control at time j, the distribution of W_j depends only on both p and n_j. Set

$$\bar{W}_j \equiv \frac{W_j - E_{\boldsymbol{\theta}_j}(W_j)|_{\boldsymbol{\theta}_j=\boldsymbol{\theta}}}{\sqrt{\mathrm{Var}_{\boldsymbol{\theta}_j}(W_j)|_{\boldsymbol{\theta}_j=\boldsymbol{\theta}}}}, \tag{2.9}$$

where both $E_{\boldsymbol{\theta}_j}(W_j)|_{\boldsymbol{\theta}_j=\boldsymbol{\theta}}$ and $\mathrm{Var}_{\boldsymbol{\theta}_j}(W_j)|_{\boldsymbol{\theta}_j=\boldsymbol{\theta}}$ are given in Appendix A.2.

2.3 Known in-control process parameters with constraint

In this subsection, assume that the in-control process parameter vector $\boldsymbol{\theta}$ is known. In practice, If the process had not been adjusted, the positive scale parameter σ_j at time j should not be smaller than in-control σ. When out of control signal occurs, σ_j should be larger than the in-control process parameter σ. Therefore, in this subsection, the process is called in control at time j if $\boldsymbol{\theta}_j = \boldsymbol{\theta}$ or out of control at time j if $\boldsymbol{\theta}_j \neq \boldsymbol{\theta}$ and $\sigma_j \geq \sigma$. The corresponding LR test statistic at time j is

$$2\left[\ell_j(\boldsymbol{\theta}_j^*) - \ell_j(\boldsymbol{\theta})\right] = 2\left[\ell_j(\hat{\boldsymbol{\theta}}_j) - \ell_j(\boldsymbol{\theta})\right] - 2\left[\ell_j(\hat{\boldsymbol{\theta}}_j) - \ell_j(\boldsymbol{\theta}_j^*)\right],$$

where $\hat{\boldsymbol{\theta}}_j$ $(= (\hat{\boldsymbol{\beta}}_j, \hat{\sigma}_j))$ is given in subsection 2.2 and $\boldsymbol{\theta}_j^*$ $(\equiv (\boldsymbol{\beta}_j^{*T}, \sigma_j^*)^T = (\hat{\boldsymbol{\beta}}_j^T, \max\{\hat{\sigma}_j, \sigma\})^T)$ is the corresponding MLE of $\boldsymbol{\theta}_j$. The corresponding score test statistic is

$$W_j^* \equiv \frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j^T} \text{Cov}_{\boldsymbol{\theta}_j}^{-1} \left(\frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \right) \frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \bigg|_{\boldsymbol{\theta}_j = \boldsymbol{\theta}} - \frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j^T} \text{Cov}_{\boldsymbol{\theta}_j}^{-1} \left(\frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \right) \frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \bigg|_{\boldsymbol{\theta}_j = \boldsymbol{\theta}_j^*}$$

$$\equiv W_j - W_j(\boldsymbol{\theta}_j^*), \tag{2.10}$$

where W_j is given in subsection 2.2 and

$$W_j(\boldsymbol{\theta}_j^*) \equiv \frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j^T} \text{Cov}_{\boldsymbol{\theta}_j}^{-1} \left(\frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \right) \frac{\partial \ell_j(\boldsymbol{\theta}_j)}{\partial \boldsymbol{\theta}_j} \bigg|_{\boldsymbol{\theta}_j = \boldsymbol{\theta}_j^*} = \frac{n_j}{2} \left[\min \left(\frac{H_{j2}}{n_j}, 1 \right) - 1 \right]^2$$

with H_{j2} being given in subsection 2.2. Then

$$W_j^* = H_{j1} + \frac{1}{2n_j}(H_{j1} + H_{j2} - n_j)^2 - \frac{n_j}{2} \left[\min \left(\frac{H_{j2}}{n_j}, 1 \right) - 1 \right]^2, \tag{2.11}$$

where H_{j1} is given in subsection 2.2. Note that the distribution of W_j^* depends only on all of p, n_j, σ_j/σ_0, and τ_j^2, where τ_j^2 is given in subsection 2.2. When the process is in control at time j, the distribution of W_j^* depends only on both p and n_j. Set

$$\bar{W}_j^* \equiv \frac{W_j^* - E_{\boldsymbol{\theta}_j}(W_j^*)|_{\boldsymbol{\theta}_j = \boldsymbol{\theta}}}{\sqrt{\text{Var}_{\boldsymbol{\theta}_j}(W_j^*)|_{\boldsymbol{\theta}_j = \boldsymbol{\theta}}}}, \tag{2.12}$$

where both $E_{\boldsymbol{\theta}_j}(W_j^*)|_{\boldsymbol{\theta}_j = \boldsymbol{\theta}}$ and $\text{Var}_{\boldsymbol{\theta}_j}(W_j^*)|_{\boldsymbol{\theta}_j = \boldsymbol{\theta}}$ are given in Appendix B.1.

2.4 Unknown in-control process parameters without constraint

In this subsection, assume that the in-control process parameter vector $\boldsymbol{\theta}$ is unknown. Assume that the historical in-control process data $\{X_0, \boldsymbol{y}_0\}$ are available. The relationship between $\boldsymbol{y}_0$ and X_0 is assumed to be

$$\boldsymbol{y}_0 = X_0\boldsymbol{\beta} + \sigma\,\varepsilon_0, \tag{2.13}$$

where $\boldsymbol{y}_0$ is an $n_0 \times 1$ response vector, X_0 is an $n_0 \times p$ known model (or design) matrix of full rank p $(< n_0)$, $\boldsymbol{\theta}$ $(\equiv (\boldsymbol{\beta}^T, \sigma)^T)$ is a $(p+1) \times 1$ unknown in-control process parameter vector, and ε_0 is a standardized error vector independent of ε_1, ε_2, $\ldots$, with $\varepsilon_0 \sim N_{n_0}(0_{n_0\times 1}, I_{n_0})$. The process is called in control at time j if $\boldsymbol{\theta}_j = \boldsymbol{\theta}$ or out of control at time j if $\boldsymbol{\theta}_j \neq \boldsymbol{\theta}$.

Then the joint p.d.f. of $(\boldsymbol{y}_0^T, \boldsymbol{y}_j^T)^T$ at time j is

$$\begin{aligned} f(\boldsymbol{y}_0, \boldsymbol{y}_j; \boldsymbol{\theta}, \boldsymbol{\theta}_j) &= f(\boldsymbol{y}_0; \boldsymbol{\theta}) \cdot f(\boldsymbol{y}_j; \boldsymbol{\theta}_j) \\ &= \frac{1}{(2\pi)^{n_0/2}\sigma^{n_0}} \exp\left\{-\frac{\|\boldsymbol{y}_0 - X_0\boldsymbol{\beta}\|^2}{2\sigma^2}\right\} \cdot f(\boldsymbol{y}_j; \boldsymbol{\theta}_j) \\ &= \frac{1}{(2\pi)^{n_0/2}\sigma^{n_0}} \exp\left\{-\frac{(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T X_0^T X_0 (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) + n_0\hat{\sigma}^2}{2\sigma^2}\right\} \cdot f(\boldsymbol{y}_j; \boldsymbol{\theta}_j), \end{aligned}$$

where $(\hat{\boldsymbol{\beta}}^T, \hat{\sigma})^T (\equiv \hat{\boldsymbol{\theta}})$ is the MLE of $\boldsymbol{\theta}$, $\hat{\boldsymbol{\beta}}$ is independent of $\hat{\sigma}$ with

$$\hat{\boldsymbol{\beta}} \;\equiv\; (X_0^T X)_0^{-1} X_0^T \boldsymbol{y}_0 \sim N_p(\boldsymbol{\beta}, \sigma^2 (X_0^T X_0)^{-1}) \tag{2.14}$$

and

$$\hat{\sigma}^2 \equiv \frac{\|\boldsymbol{y}_0 - X_0\hat{\boldsymbol{\beta}}\|^2}{n_0} = \frac{\left\|\left[I_{n_j} - X_j(X_0^T X_0)^{-1} X_0^T\right]\boldsymbol{y}_0\right\|^2}{n_0} \sim \frac{\sigma^2}{n_0}\chi^2_{n_0-p}. \tag{2.15}$$

The log-likelihood function for $(\boldsymbol{\theta}^T, \boldsymbol{\theta}_j^T)^T$ at time j is

$$\begin{aligned}\ell_{0,j}(\boldsymbol{\theta}, \boldsymbol{\theta}_j) &\equiv \log\left[f(\boldsymbol{y}_0; \boldsymbol{\theta}) \cdot f(\boldsymbol{y}_j; \boldsymbol{\theta}_j)\right] = \log\left[f(\boldsymbol{y}_0; \boldsymbol{\theta})\right] + \log\left[f(\boldsymbol{y}_j; \boldsymbol{\theta}_j)\right] \\ &= -\frac{n_0}{2}\log(2\pi) - \frac{n_0}{2}\log(\sigma^2) - \frac{\|\boldsymbol{y}_0 - X_0\boldsymbol{\beta}\|^2}{2\sigma^2} + \ell_j(\boldsymbol{\theta}_j).\end{aligned}$$

The corresponding LR and score test statistics at time j are

$$2\left[\ell_{0,j}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}_j) - \ell_{0,j}(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}_j)\right]$$

and

$$\frac{\partial \ell_{0,j}(\boldsymbol{\theta}, \boldsymbol{\theta}_j)}{\partial(\boldsymbol{\theta}^T, \boldsymbol{\theta}_j^T)} \mathrm{Cov}^{-1}_{\boldsymbol{\theta}, \boldsymbol{\theta}_j}\left(\frac{\partial \ell_{0,j}(\boldsymbol{\theta}, \boldsymbol{\theta}_j)}{\partial(\boldsymbol{\theta}, \boldsymbol{\theta}_j)}\right) \frac{\partial \ell_{0,j}(\boldsymbol{\theta}, \boldsymbol{\theta}_j)}{\partial(\boldsymbol{\theta}^T, \boldsymbol{\theta}_j^T)^T}\Bigg|_{\boldsymbol{\theta}=\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}_j=\tilde{\boldsymbol{\theta}}_j} (\equiv W_{0,j}),$$

where $(\tilde{\boldsymbol{\theta}}^T, \tilde{\boldsymbol{\theta}}_j^T)^T$ are the MLE of $(\boldsymbol{\theta}^T, \boldsymbol{\theta}_j^T)^T$ when the process is in control at time j with

$$\begin{aligned}\tilde{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}_j &= (X_0^T X_0 + X_j^T X_j)^{-1}(X_0^T \boldsymbol{y}_0 + X_j^T \boldsymbol{y}_j) \\ &= \left(X_0^T X_0 + X_j^T X_j\right)^{-1}\left(X_0^T X_0\hat{\boldsymbol{\beta}} + X_j^T X_j \hat{\boldsymbol{\beta}}_j\right) \\ \tilde{\sigma}_0 = \tilde{\sigma}_j &= \sqrt{\frac{\|\boldsymbol{y}_0 - X_0\tilde{\boldsymbol{\beta}}\|^2 + \|\boldsymbol{y}_j - X_j\tilde{\boldsymbol{\beta}}_j\|^2}{n_0 + n_j}} \\ &= \sqrt{\frac{n_0\hat{\sigma}^2 + n_j\hat{\sigma}_j^2 + (\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}})^T X_0^T X_0(\hat{\boldsymbol{\beta}} - \tilde{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}}_j - \tilde{\boldsymbol{\beta}})^T X_j^T X_j(\hat{\boldsymbol{\beta}}_j - \tilde{\boldsymbol{\beta}})}{n_0 + n_j}} \\ &= \sqrt{\frac{n_0\hat{\sigma}^2 + n_j\hat{\sigma}_j^2 + (\hat{\boldsymbol{\beta}}_j - \hat{\boldsymbol{\beta}})^T[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}]^{-1}(\hat{\boldsymbol{\beta}}_j - \hat{\boldsymbol{\beta}})}{n_0 + n_j}};\end{aligned} \tag{2.16}$$

and $\partial\ell_{0,j}(\boldsymbol{\theta},\boldsymbol{\theta}_j)/\partial(\boldsymbol{\theta}^T,\boldsymbol{\theta}_j^T)^T$ is the score function for $(\boldsymbol{\theta}^T,\boldsymbol{\theta}_j^T)^T$ at time j such that

$$
\begin{aligned}
\frac{\partial\ell_{0,j}(\boldsymbol{\theta},\boldsymbol{\theta}_j)}{\partial\boldsymbol{\beta}} &= \frac{X_0^T(\boldsymbol{y}_0 - X_0\boldsymbol{\beta})}{\sigma^2} \sim N_p\left(0_{p\times 1}, \frac{X_0^T X_0}{\sigma^2}\right), \\
\frac{\partial\ell_{0,j}(\boldsymbol{\theta},\boldsymbol{\theta}_j)}{\partial\sigma} &= \frac{\|\boldsymbol{y}_0 - X_0\boldsymbol{\beta}\|^2}{\sigma^3} - \frac{n_0}{\sigma} \sim \frac{\chi^2_{n_0} - n_0}{\sigma}, \\
\frac{\partial\ell_{0,j}(\boldsymbol{\theta},\boldsymbol{\theta}_j)}{\partial\boldsymbol{\beta}_j} &= \frac{X_j^T(\boldsymbol{y}_j - X_j\boldsymbol{\beta}_j)}{\sigma_j^2} \sim N_p\left(0_{p\times 1}, \frac{X_j^T X_j}{\sigma_j^2}\right), \\
\frac{\partial\ell_{0,j}(\boldsymbol{\theta},\boldsymbol{\theta}_j)}{\partial\sigma_j} &= \frac{\|\boldsymbol{y}_j - X_j\boldsymbol{\beta}_j\|^2}{\sigma_j^3} - \frac{n_j}{\sigma_j} \sim \frac{\chi^2_{n_j} - n_j}{\sigma_j}.
\end{aligned}
$$

When the process is in control at time j,

$$
2\left[\ell_{0,j}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{\theta}}_j) - \ell_{0,j}(\tilde{\boldsymbol{\theta}},\tilde{\boldsymbol{\theta}}_j)\right] = W_{0,j} + O_p\left(\frac{1}{\sqrt{min\{n_0, n_j\}}}\right), \qquad (2.17)
$$

where

$$
\begin{aligned}
& W_{0,j} \\
=& \frac{(\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\beta}})^T X_0^T X_0(\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\beta}})}{\tilde{\sigma}^2}+\frac{1}{2n_0}\left[\frac{(\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\beta}})^T X_0^T X_0(\hat{\boldsymbol{\beta}}-\tilde{\boldsymbol{\beta}})}{\tilde{\sigma}^2}+\frac{n_0\hat{\sigma}^2}{\tilde{\sigma}^2}-n_0\right]^2 \\
&+\frac{(\hat{\boldsymbol{\beta}}_j-\tilde{\boldsymbol{\beta}}_j)^T X_j^T X_j(\hat{\boldsymbol{\beta}}_j-\tilde{\boldsymbol{\beta}}_j)}{\tilde{\sigma}_j^2}+\frac{1}{2n_j}\left[\frac{(\hat{\boldsymbol{\beta}}_j-\tilde{\boldsymbol{\beta}}_j)^T X_j^T X_j(\hat{\boldsymbol{\beta}}_j-\tilde{\boldsymbol{\beta}}_j)}{\tilde{\sigma}_j^2}+\frac{n_j\hat{\sigma}_j^2}{\tilde{\sigma}_j^2}-n_j\right]^2 \\
=& \frac{(n_0+n_j)(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})^T A_{0j}(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})}{n_0\hat{\sigma}^2+n_j\hat{\sigma}_j^2+(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})^T A_{0j}(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})} \\
&+\frac{1}{2n_0}\left\{\frac{(n_0+n_j)\left[(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})^T X_j^T X_j B_{0j} X_0^T X_0 B_{0j} X_j^T X_j(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})+n_0\hat{\sigma}^2\right]}{n_0\hat{\sigma}^2+n_j\hat{\sigma}_j^2+(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})^T A_{0j}(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})}-n_0\right\}^2 \\
&+\frac{1}{2n_j}\left\{\frac{(n_0+n_j)\left[(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})^T X_0^T X_0 B_{0j} X_j^T X_j B_{0j} X_0^T X_0(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})+n_j\hat{\sigma}_j^2\right]}{n_0\hat{\sigma}^2+n_j\hat{\sigma}_j^2+(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})^T A_{0j}(\hat{\boldsymbol{\beta}}_j-\hat{\boldsymbol{\beta}})}-n_j\right\}^2
\end{aligned}
$$

with

$$A_{0j} \equiv [(X_0^T X_0)^{-1}+(X_j^T X_j)^{-1}]^{-1}$$

and

$$B_{0j} \equiv (X_0^T X_0+X_j^T X_j)^{-1}.$$

When the process is in control at time j, the distribution of $W_{0,j}$ depends only on all of p, n_0, n_j and p eigenvalues of $(X_0^T X_0)^{-1/2} X_j^T X_j (X_0^T X_0)^{-1/2}$. Set

$$\bar{W}_{0,j} \equiv \frac{W_{0,j} - E_{\boldsymbol{\theta},\boldsymbol{\theta}_j}(W_{0,j})|_{\theta=\tilde{\theta},\theta_j=\tilde{\theta}_j}}{\sqrt{\mathrm{Var}_{\boldsymbol{\theta},\boldsymbol{\theta}_j}(W_{0,j})|_{\theta=\tilde{\theta},\theta_j=\tilde{\theta}_j}}}, \tag{2.18}$$

where both $E_{\boldsymbol{\theta},\boldsymbol{\theta}_j}(W_{0,j})|_{\theta=\tilde{\theta},\theta_j=\tilde{\theta}_j}$ and $\mathrm{Var}_{\boldsymbol{\theta},\boldsymbol{\theta}_j}(W_{0,j})|_{\theta=\tilde{\theta},\theta_j=\tilde{\theta}_j}$ do not depend on the unknown in-control process parameter vector $\boldsymbol{\theta}$ and are given in Appendix C.1.

2.5 Unknown in-control process parameters with constraint

In this subsection, assume that the in-control process parameter vector $\boldsymbol{\theta}$ is unknown. And the positive scale parameter σ_j at time j $(j \geq 1)$ is larger than or equal to the in-control process parameter σ if the process has not been adjusted until time j. Therefore, in this subsection, the process is called in control at time j if $\boldsymbol{\theta}_j = \boldsymbol{\theta}$ or out of control at time j if $\boldsymbol{\theta}_j \neq \boldsymbol{\theta}$ and $\sigma_j \geq \sigma$. The corresponding LR test statistic at time j is

$$2\left[\ell_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*) - \ell_{0,j}(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}_j)\right] = 2\left[\ell_{0,j}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}_j) - \ell_{0,j}(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\theta}}_j)\right] - 2\left[\ell_{0,j}(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\theta}}_j) - \ell_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*)\right]$$

where $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{\theta}}_j$, $\tilde{\boldsymbol{\theta}}$, and $\tilde{\boldsymbol{\theta}}_j$ are given in subsection 2.3, $(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*)$ $(\equiv (\boldsymbol{\beta}^*,\ \sigma^*,\ \boldsymbol{\beta}_j^*,\ \sigma_j^*))$ are the MLE of $(\boldsymbol{\theta}, \boldsymbol{\theta}_j)$ under out-of-control status with

$$\boldsymbol{\beta}^* = \hat{\boldsymbol{\beta}},\ \sigma^* = \hat{\sigma} \cdot 1_{\{\hat{\sigma}_j \geq \hat{\sigma}\}} + \sqrt{\frac{\|\boldsymbol{y}_0 - X_0\hat{\boldsymbol{\beta}}\|^2 + \|\boldsymbol{y}_j - X_j\hat{\boldsymbol{\beta}}_j\|^2}{n_0 + n_j}} \cdot 1_{\{\hat{\sigma}_j < \hat{\sigma}\}},$$

and

$$\boldsymbol{\beta}_j^* = \hat{\boldsymbol{\beta}}_j,\ \sigma_j^* = \hat{\sigma}_j \cdot 1_{\{\hat{\sigma}_j \geq \hat{\sigma}\}} + \sqrt{\frac{\|\boldsymbol{y}_0 - X_0\hat{\boldsymbol{\beta}}\|^2 + \|\boldsymbol{y}_j - X_j\hat{\boldsymbol{\beta}}_j\|^2}{n_0 + n_j}} \cdot 1_{\{\hat{\sigma}_j < \hat{\sigma}\}}.$$

The corresponding score test statistic is

$$W_{0,j}^* \equiv W_{0,j} - W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*),$$

where $W_{0,j}$ is given in subsection 2.4, and

$$W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*) \equiv \frac{\partial \ell_{0,j}(\boldsymbol{\theta}, \boldsymbol{\theta}_j)}{\partial(\boldsymbol{\theta}^T, \boldsymbol{\theta}_j^T)} \mathrm{Cov}_{\boldsymbol{\theta},\boldsymbol{\theta}_j}^{-1} \left(\frac{\partial \ell_{0,j}(\boldsymbol{\theta}, \boldsymbol{\theta}_j)}{\partial(\boldsymbol{\theta}, \boldsymbol{\theta}_j)} \right) \frac{\partial \ell_{0,j}(\boldsymbol{\theta}, \boldsymbol{\theta}_j)}{\partial(\boldsymbol{\theta}^T, \boldsymbol{\theta}_j^T)^T} \bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*, \boldsymbol{\theta}_j=\boldsymbol{\theta}_j^*}$$

$$= \frac{n_0}{2}\left(\frac{\hat{\sigma}^2}{\sigma^{*2}} - 1\right)^2 + \frac{n_j}{2}\left(\frac{\hat{\sigma}_j^2}{\sigma_j^{*2}} - 1\right)^2 = (n_0 + n_j)\left(\frac{n_0 n_j}{2}\right)\left(\frac{\hat{\sigma}^2 - \hat{\sigma}_j^2}{n_0\hat{\sigma}^2 + n_j\hat{\sigma}_j^2}\right)^2 \cdot 1_{\{\hat{\sigma}_j < \hat{\sigma}\}}. \tag{2.19}$$

When the process is in control at time j, the distribution of $W_{0,j}^*$ depends only on p, n_0, and n_j. Set

$$\bar{W}_{0,j}^* \equiv \frac{W_{0,j}^* - E_{\boldsymbol{\theta},\boldsymbol{\theta}_j}(W_{0,j}^*)|_{\theta=\tilde{\theta},\theta_j=\tilde{\theta}_j}}{\sqrt{\mathrm{Var}_{\boldsymbol{\theta},\boldsymbol{\theta}_j}(W_{0,j}^*)|_{\theta=\tilde{\theta},\theta_j=\tilde{\theta}_j}}}, \tag{2.20}$$

where both $E_{\boldsymbol{\theta},\boldsymbol{\theta}_j}(W_{0,j}^*)|_{\theta=\tilde{\theta},\theta_j=\tilde{\theta}_j}$ and $\mathrm{Var}_{\boldsymbol{\theta},\boldsymbol{\theta}_j}(W_{0,j}^*)|_{\theta=\tilde{\theta},\theta_j=\tilde{\theta}_j}$ are given in Appendix D.1.

2.6 Proposed monitoring scheme

The EWMA sequence is defined as

$$\begin{cases} U_0 \equiv 0, \\ U_j \equiv (1-\lambda)U_{j-1} + \lambda \tilde{W}_j, \quad j = 1, 2, \cdots, \end{cases} \tag{2.21}$$

where λ is a smoothing parameter in $(0, 1]$, and $\tilde{W}_j = \bar{W}_j$ in subsection 2.2, or $\tilde{W}_j = \bar{W}^*_j$ in subsection 2.3, $\tilde{W}_j = \bar{W}_{0,j}$ in subsection 2.4, $\tilde{W}_j = \bar{W}^*_{0,j}$ in subsection 2.5. Then, standardize the statistic and define the control chart statistic as

$$U^*_j \equiv \frac{U_j - E_{\boldsymbol{\theta}}(U_j)}{\sqrt{\text{Var}_{\boldsymbol{\theta}}(U_j)}}, \tag{2.22}$$

where

$$E_{\boldsymbol{\theta}}(U_j) = 0,$$

and

$$\text{Var}_{\boldsymbol{\theta}}(U_j) = \frac{\lambda\left[1 - (1-\lambda)^{2j}\right]}{2-\lambda},$$

for known $\boldsymbol{\theta}$ case;

$$E_{\boldsymbol{\theta}}(U_j) = 0,$$

and

$$\text{Var}_{\boldsymbol{\theta}}(U_j) = \frac{\lambda\left[1 - (1-\lambda)^{2j}\right]}{2-\lambda} + 2\lambda^2 \sum_{0 \le k_1 < k_2 \le j-1} (1-\lambda)^{k_1+k_2} \text{Cov}(\tilde{W}_{j-k_1}, \tilde{W}_{j-k_2}),$$

for unknown $\boldsymbol{\theta}$ case. Because $\tilde{W}_{j-k_1}$ and $\tilde{W}_{j-k_2}$ are correlative, so the calculation of $\text{Cov}(\tilde{W}_{j-k_1}, \tilde{W}_{j-k_2})$ will be very complex when X_js are not design matrix. Therefore, here recommend X_0 and X_j as

$$X_0 = \begin{pmatrix} X \\ \vdots \\ X \end{pmatrix}_{(mn)\times p}, X_j = \begin{pmatrix} X \\ \vdots \\ X \end{pmatrix}_{(m_j n)\times p}, j \geq 1,$$

where X is a $n \times p$ design matrix, then $\mathrm{Cov}(\tilde{W}_{j-k_1}, \tilde{W}_{j-k_2})$ will only depend on $n,\ p,\ m,\ m_{j-k_1},\ m_{j-k_1}$

Out-of-control signal occurs at time j if

$$U_j^* > C,$$

where $C > 0$ is chosen by a specified in-control ARL ($\equiv$ ARL$_0$). C only depend on ARL$_0$, λ, p, n_1, n_2, $\ldots$ for known $\boldsymbol{\theta}$ case and depend on ARL$_0$, λ, p, n, m, m_1, m_2, $\ldots$ for unknown $\boldsymbol{\theta}$ case.

3 A Simulation Study

In this Section, we explain how to simulate the proposed EWMA control chart for comparison with the Kim *et al.* (2003) and Zou *et al.* (2007). Consider the case of constrained EWMA control chart with known parameters first. If the process is in control, in equation (2.11), the H_{j1} follows chi-squared distribution with p degrees of freedom, H_{j2} follows chi-squared distribution with $n_j - p$ degrees. Use this property to generate test statistic and find control limit C.

Step 1 : Choose the specific ARL_0 and smoothing parameter λ, here given ARL_0=200 and λ=0.2 and 50,000 simulations, The same assumptions with Kim *et al.* (2003) and Zou *et al.* (2006)

Step 2 : Generate 200 H_{j1} and H_{j2}, and calculated 200 W_j^*, U_j, and U_j^*, for $j = 1, 2, \ldots, 200$

Step 3 : $c \equiv \max\{U_1^*, U_2^*, \ldots, U_{200}^*\}$.

Step 4 : Repeat Step (2) $\sim$ Step (3) 50,000 times, obtain 50,000 c, and make them to sort $(c_{(1)} < c_{(2)} < \ldots < c_{(50000)})$.

Step 5 : Choose the median of the 50,000 c as the control limit. Use this control limit, make 50,000 time simulation, then obtain 50,000 Run Length, and compute the ARL.

Step 6 : Use bisection method if the ARL $>$ 200 then make control limit as the smaller c in 50,000 c, if the ARL $<$ 200 then make control limit as the bigger c in 50,000 c.

Step 7 : Repeat step (5) $\sim$ (6), until the 199.5<ARL<200.5, then obtain-control limit C.

It can be easily generate the out-of-control W_j^* by equation (2.11), and use the control limit which obtained previously to calculate the out-of control ARL (ARL_1) with 50,000 simulations.

Next, consider the case of constrained EWMA control chart with unknown parameters. The procedure is same as case of constrained EWMA control chart with known parameters except the generation of U^*. In subsection 2.5, $W_{0,j}^*$ consists of $\hat{\boldsymbol{\beta}}$, $\hat{\boldsymbol{\beta}}_j$, $\hat{\sigma}$, and $\hat{\sigma}_j$, where

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &\sim N_p(\boldsymbol{\beta}, \sigma^2(X_0^T X_0)^{-1}), \\
\hat{\boldsymbol{\beta}}_j &\sim N_p(\boldsymbol{\beta}_j, \sigma_j^2(X_j^T X_j)^{-1}), \\
\hat{\sigma} &\sim \sigma^2/n_0 \chi^2_{n_0-p,} \\
\hat{\sigma}_j &\sim \sigma_j^2/n_j \chi^2_{n_j-p,}
\end{aligned}$$

and they are independent respectively. $W_{0,j}^*$ and $W_{0,j'}^*$ are depends on $p, n_0, n_1, \ldots (j \neq j')$, and therefore the X_0 and X_j must be design matrix. Then generate U^* through above property.

We consider the case from Kim *et al.* (2003) and Zou *et al.* (2007), the simplest case of model (2.2) with $p = 2$, in-control parameters $\boldsymbol{\beta} \equiv (\beta_0, \beta_1)^T = (3, 2)^T$, $\sigma = 1$, fixed

X_j,

$$X_j \equiv X = \begin{pmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 6 \\ 1 & 8 \end{pmatrix}, j \geq 1,$$

and

$$X_0 = \begin{pmatrix} X \\ \vdots \\ X \end{pmatrix}_{4m \times p}.$$

4 Conclusions

4.1 Monitoring performance comparisons

Compared proposed EWMA chart with the KMW, ZTW and Huang (2012) EWMA chart by out of control ARL (ARL_1), the ARL_1 of three control chart with known parameters are given in Table 5.1∼5.3. The process parameter β_0 is changed to $\beta_0 + \delta_0\sigma$, and β_1 is changed to $\beta_1 + \delta_1\sigma$ in Table 5.1 and Table 5.2. The ARL_1 of different ratio with σ and σ_j are presented in Table 5.3.

First focus on the KMW and ZTW, In Table 5.1 and Table 5.2. The proposed EWMA chart is favorable for detecting large shift in β_0 and β_1, but when shift is moderate or small, proposed EWMA chart has a significant adverse, it maybe cause by the property of score test and LR test. In the case of large shift or n_j, the performance should be good, on the other hand, the case of small shift and n_j, have the opposite result. The n_j is fixed to 4, and $n_j - p = 2$, therefore, this result is not surprising.

In the table 5.3. show that performance in all change case of the proposed EWMA chart are superior to others chart. In equation (2.7) and (2.8) show that changes of β_0 and β_1 only affect the H_{j1}, but changes of σ^2 affect not only H_{j1} but also H_{j2}, therefore, the proposed EWMA chart is more sensitive when σ^2 shifted.

Table 5.4 presented ARL_1 with process parameters $\boldsymbol{\beta}$ and σ change simultaneously. And the unknown in-control process parameters case also presented. When the $m = 125$, ARL_1 performance will be very close the known in-control process parameters case.

4.2 Modify monitoring scheme (1)

In Huang (2012), he proposed an EWMA control chat based on LR test statistic, it is similar with the method we proposed. It should be similar results in theory, but according to Table 5.1∼5.3, they are a little difference. The HEWMA is better than proposed EWMA when the $\boldsymbol{\beta}$ shifted, and it is worse than proposed EWMA when the σ shifted.

Here proposed an improved method, we transform the score statistic to make it more sensitive for the detection of $\boldsymbol{\beta}$.

$$\begin{aligned} W_j &= \frac{(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})^T X_j^T X_j (\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})}{\sigma_0^2} \\ &\quad + a \left[\Phi^{-1} \left(F_{\chi^2_{n_j}} \left(\frac{(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})^T X_j^T X_j (\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})}{\sigma^2} + \frac{n_j \hat{\sigma}_j^2}{\sigma^2} \right) \right) \right]^2, \end{aligned}$$

for two-sided case, and

$$\begin{aligned} W_j^* &= \frac{(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})^T X_j^T X_j (\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})}{\sigma^2} \\ &\quad + a F^{-1}_{\chi^2_1} \left(F_{\chi^2_{n_j}} \left(\frac{(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})^T X_j^T X_j (\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})}{\sigma^2} + \frac{n_j \hat{\sigma}_j^2}{\sigma^2} \right) \right), \end{aligned}$$

for one-sided case, where $a \in (0, \infty)$. When the process is in control at time j, the distribution of W_j and W_j^* depend only on p and n_j. The greater a is, and the more sensitive for the detection of σ. On the other hand, here needing to a more sensitive detection of $\boldsymbol{\beta}$, and therefore the ARL_1 performance with $0 < a < 1$ are given in Table 5.6 and Table 5.7. When a the smaller, the performance will be almost the same as with the HEWMA or better. We found very bad performance for the detection of σ shifted in table 5.8. When the emphasis on the sift of $\boldsymbol{\beta}$, we recommend the use this method, and when the emphasis on the shift of σ, we recommend using the methods mentioned earlier.

4.3 Modify monitoring scheme (2)

In the previous subsection, even the use of the modified statistic $\boldsymbol{\beta}$ shifted ARL_1 performance still worse than KMW and ZTW. And think it is $(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})^T X_j^T X_j (\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta})/\sigma^2$ terms caused. It makes all the factors of $\boldsymbol{\beta}$ changes into a single value. Then propose a improve method. First use the MEWMA Chart. Let

$$
\begin{aligned}
W_0 &\equiv 0_{p+1\times 1} \\
W_j &\equiv \begin{pmatrix} W_{j1} \\ \vdots \\ W_{j,p+1} \end{pmatrix} = \lambda \begin{pmatrix} \partial\ell_j/\partial\boldsymbol{\beta}_j|_{\theta_j=\theta} \\ \Phi^{-1}\left(F_{\partial\ell_j/\partial\sigma_j|_{\theta_j=\theta}}\left(\partial\ell_j/\partial\sigma_j|_{\theta_j=\theta};\boldsymbol{\theta}\right)\right) \end{pmatrix} + (1-\lambda)W_{j-1},\ j=1,2,3,\cdots, \\
&= \lambda \begin{pmatrix} X_j^T(\boldsymbol{y}_j - X_j\boldsymbol{\beta})/\sigma^2 \\ \Phi^{-1}\left(F_{\chi^2_{n_j}}\left(\|\boldsymbol{y}_j - X_j\boldsymbol{\beta}\|^2/\sigma^2\right)\right) \end{pmatrix} + (1-\lambda)W_{j-1},\ j=1,2,3,\cdots,
\end{aligned}
\tag{4.1}
$$

where $X_j^T(\boldsymbol{y}_j - X_j\boldsymbol{\beta})/\sigma^2$ and $\Phi^{-1}\left(F_{\chi^2_{n_j}}\left(\|\boldsymbol{y}_j - X_j\boldsymbol{\beta}\|^2/\sigma^2\right)\right)$ are not independent but

$$\mathrm{Cov}_{\theta_j=\theta}\left(X_j^T(\boldsymbol{y}_j - X_j\boldsymbol{\beta})/\sigma^2, \Phi^{-1}\left(F_{\chi^2_{n_j}}\left(\|\boldsymbol{y}_j - X_j\boldsymbol{\beta}\|^2/\sigma^2\right)\right)\right) = 0_{p\times 1},$$

and

$$\mathrm{Cov}_{\theta_1,\ldots,\theta_j}(W_j)\big|_{\theta_1=\ldots=\theta_j=\theta} = \lambda^2\sum_{k=1}^{j}(1-\lambda)^{2(j-k)}\begin{pmatrix} X_k^TX_k/\sigma^2 & 0_{p\times 1} \\ 0_{p\times 1}^T & 1 \end{pmatrix},\ j = 1, 2, \ldots.$$

Then make the two-sided score test statistic as

$$U_j = \begin{pmatrix} W_{j1} \\ \vdots \\ W_{jp} \end{pmatrix}^T \mathrm{Cov}^{-1}\begin{pmatrix} W_{j1} \\ \vdots \\ W_{jp} \end{pmatrix}\Bigg|_{\theta_1=\ldots=\theta_j=\theta}\begin{pmatrix} W_{j1} \\ \vdots \\ W_{jp} \end{pmatrix} + a\frac{W_{j,p+1}^2}{\lambda^2\sum_{k=1}^{j}(1-\lambda)^{2(j-k)}}, \tag{4.2}$$

and the one-sided score test statistic as

$$U_j^* = \begin{pmatrix} W_{j1} \\ \vdots \\ W_{jp} \end{pmatrix}^T \mathrm{Cov}^{-1}\begin{pmatrix} W_{j1} \\ \vdots \\ W_{jp} \end{pmatrix}\Bigg|_{\theta_1=\ldots=\theta_j=\theta}\begin{pmatrix} W_{j1} \\ \vdots \\ W_{jp} \end{pmatrix} + aF_{\chi^2_1}^{-1}\left(\Phi\left(\frac{W_{j,p+1}}{\lambda\sqrt{\sum_{k=1}^{j}(1-\lambda)^{2(j-k)}}}\right)\right), \tag{4.3}$$

where $a \in (0,\ \infty)$. Similar with before, the greater a is, and the more sensitive for the detection of σ, whereas other is sensitive for the detection of $\boldsymbol{\beta}$. The simulation results are presented in Table 5.9$\sim$5.11.

In table 5.9 and table 5.10 shows the simulation results of the $\boldsymbol{\beta}$ shifted with two-sided case. The ARL_1 performance of the improve method is very close with ZTW and KMW, and when the constant a is getting smaller even better than their. In table 5.11 shows the simulation results of the σ shifted with two-sided and one-sided case. When constant $a = 1$, one-sided case more sensitive than two-sided, and the they are better than ZTW and KMW, but slightly worse than HEMWA. And the greater constant a is, the worse performance becomes. This improve method can be seen here in the case of whether $\boldsymbol{\beta}$ shifted or σ shifted has good, and to focus on detection of shifts in the $\boldsymbol{\beta}$ or σ by by adjusting constant a.

5 Future Work

In this paper, the MEWMA performance seems better than the performance of EWMA based on score test statistics. In future work, it can focus on MEWMA. Consider the nonlinear model. e.g., $\boldsymbol{y}_j = u_j(X_j; \boldsymbol{\beta}_j) + \sigma_j \varepsilon_j$, where $u_j(\cdot\ ;\ \cdot)$ is a known function for $j = 1, 2, \cdots$. May also consider the error term is not a normal distribution. e.g., t-distribution. And considered the MEWMA method based on score test statistics will have a good performance.

References

[1] Castagliola, P., Celano, G., and Fichera, S. (2006) Monitoring process variability using EWMA. *Springer Handbook of Engineering Statistics*, 291-325. Ch.17.

[2] Huang, W-C. (2012) An exponentially weighted moving average control chart based on likelihood ratio test statistics for monitoring general linear profiles. Master thesis, National Chiao Tung University, Hsinchu, Taiwan.

[3] Jensen, W. A., Champ, C. W., and Woodall, W. H. (2006) Effects of parameter estimation on control chart properties: a literature review. *Journal of Quality Technology*, **38**, 156-167.

[4] Jones, L. A., Champ, C. W., and Rigdon, S. E. (2001) The performance of exponentially weighted moving average charts with estimated parameters. *Technometrics*, **43**, 156-167.

[5] Kim, K., Mahmoud, M., and Woodall, W. H. (2003) On the monitoring of linear profiles. *Journal of Quality Technology*, **35**, 317-328.

[6] Kots, Balakrishnan and Johnson. (2000) Continuous multivariate distributions volume 1 second edition.

[7] Page, E. S. (1954) Continuous inspection schemes. *Biometrika*, **41**, 100-115.

[8] Roberts, S. W. (1959) Control chart tests based on geometric moving averages. *Technometrics*, **1**, 239-250.

[9] Shewhart, W. A. (1931) *Economic control of quality of manufactured product*, American Society for Quality, Vol. 509.

[10] Zou, C., Lin, Y., and Wang, Z. (2009) Comparisons of control schemes for monitoring the means of processes subject to drifts. *Metrika*, **70**, 141-163.

[11] Zou, C., Tsung, F., and Wang, Z. (2007) Monitoring general linear profiles using multivariate exponentially weighted moving average schemes. *Technometrics*, **49**, 395-408.

[12] Zou, C., Zhang, Y., and Wang, Z. (2006) A control chart based on a change-point model for monitoring linear profiles. *Transactions*, **38**, 1093-1103.

[13] Zou, C., Zhang, J., and Wang, Z. (2010) A control chart based on likelihood ratio test for monitoring process mean and variability. *Quality and Reliability Engineering International*, **26**, 63-73.

Table 5.1: Out of control ARL comparisons between EWMA with known in-control parameters and constraint case, ZTW, KMW and HEWMA charts for shifts in β_0, where $\boldsymbol{\theta}_j = (\beta_0 + \delta_0\sigma, \beta_1, \sigma)^T$ and $\tau_j^2 = 4\delta_0^2$ for $j \geq 1$.

		ARL_1			
δ_0	τ_j^2	EWMA	ZTW	KMW	HEWMA
0.1000	0.0400	178.7	131.5	133.7	171.1
0.2000	0.1600	133.1	59.9	59.1	113.1
0.3000	0.3600	86.7	29.6	28.3	67.2
0.4000	0.6400	53.4	17.2	16.2	38.1
0.5000	1.0000	32.9	11.5	10.7	22.1
0.6000	1.4400	20.5	8.5	7.9	13.8
0.8000	2.5600	9.1	5.8	5.1	6.1
1.0000	4.0000	4.8	4.1	3.8	3.4
1.5000	9.0000	1.7	2.6	2.4	1.5
2.0000	16.000	1.2	2.0	1.9	1.1

Table 5.2: Out of control ARL comparisons between EWMA with known in-control parameters and constraint case, ZTW, KMW and HEWMA charts for shifts in β_1, where $\boldsymbol{\theta}_j = (\beta_0, \beta_1 + \delta_1\sigma, \sigma)^T$ and $\tau_j^2 = 120\delta_1^2$ for $j \geq 1$.

		ARL_1			
δ_1	τ_j^2	EWMA	ZTW	KMW	HEWMA
0.0250	0.0750	162.4	99.0	101.6	153.0
0.0375	0.1688	130.1	57.4	61.0	112.3
0.0500	0.3000	97.0	35.0	36.5	77.4
0.0625	0.4688	70.9	23.1	24.6	53.2
0.0750	0.6750	51.1	16.4	17.0	35.4
0.1000	1.2000	26.2	9.8	10.3	17.4
0.1250	1.8750	14.2	6.9	7.2	9.5
0.1500	2.7000	8.4	5.3	5.5	5.8
0.2000	4.8000	3.6	3.7	3.8	2.8
0.2500	7.5000	2.2	2.9	2.9	1.8

Table 5.3: Out of control ARL comparisons between EWMA with known in-control parameters and constraint case , ZTW, KMW and HEWMA charts for shifts in σ, where $\boldsymbol{\theta}_j = (\beta_0, \beta_1, \delta\sigma)^T$ for $j \geq 1$.

	ARL_1			
δ	EWMA	ZTW	KMW	HEWMA
1.10	49.0	76.2	72.8	51.1
1.15	29.4	48.7	48.1	30.2
1.20	19.0	33.2	33.5	20.0
1.25	13.4	24.1	24.9	14.6
1.30	10.0	18.4	19.4	11.0
1.40	6.4	12.1	12.7	7.0
1.60	3.8	7.0	7.2	4.0
1.80	2.4	4.9	5.1	2.8
2.20	1.6	3.1	3.2	1.8
2.60	1.3	2.3	2.5	1.4
3.00	1.2	1.9	2.1	1.3

Table 5.4: ARLs with unknown in-control parameters and constraint case for shifts in $\boldsymbol{\theta}$ with $\boldsymbol{\theta}_j = (\boldsymbol{\beta}_j^T, \delta\sigma)^T$ and $\tau_j^2 = \tau^2/\delta^2$ for $j \geq 1$

	$m = 5$ $m = 25$ $m = 125$ $m = \infty$	δ 1.1^0	1.1^1	1.1^2	1.1^3	1.1^4	1.1^5	1.1^6	1.1^7	1.1^8	1.1^9	1.1^{10}	1.1^{11}	1.1^{12}
	0.0^2	200.0	71.6	31.9	14.9	9.5	5.3	4.1	3.0	2.1	1.9	1.7	1.5	1.4
		200.0	54.9	20.2	9.9	6.0	3.9	2.9	2.3	1.9	1.5	1.4	1.3	1.2
		200.0	50.1	18.7	9.4	5.4	3.5	2.7	2.0	1.8	1.4	1.3	1.3	1.2
		200.0	49.0	18.0	8.9	4.8	3.3	2.5	1.9	1.7	1.3	1.3	1.2	1.2
	0.2^2	189.2	69.9	28.5	13.7	9.2	5.1	3.9	2.9	2.0	2.0	1.7	1.6	1.4
		184.3	52.0	19.8	9.3	5.6	3.7	2.8	2.3	1.8	1.5	1.4	1.3	1.2
		182.2	47.1	18.0	8.9	5.3	3.2	2.5	2.0	1.7	1.3	1.3	1.2	1.2
		178.7	44.2	17.0	8.4	4.1	3.0	2.2	1.6	1.4	1.3	1.3	1.2	1.2
	0.4^2	150.1	62.1	27.3	12.9	8.8	4.9	3.7	2.7	2.0	1.9	1.6	1.5	1.4
		140.1	54.3	18.0	9.1	5.4	3.5	2.7	2.2	1.7	1.4	1.3	1.3	1.2
		136.8	49.2	17.0	8.5	4.8	3.1	2.5	2.0	1.7	1.3	1.3	1.2	1.2
		133.1	40.1	14.1	7.9	3.9	2.6	2.0	1.5	1.4	1.3	1.3	1.2	1.2
	0.6^2	99.9	45.2	22.7	10.9	8.3	4.8	3.5	2.4	2.0	2.0	1.5	1.5	1.4
		93.7	37.8	15.0	8.4	5.3	3.3	2.6	2.0	1.7	1.4	1.3	1.2	1.2
		89.1	33.5	14.1	7.5	4.6	2.8	2.2	1.8	1.6	1.3	1.3	1.2	1.2
		86.7	30.8	13.1	7.1	3.4	2.5	1.9	1.5	1.4	1.3	1.3	1.2	1.2
	0.8^2	70.2	43.1	20.9	10.1	7.1	4.1	3.3	2.2	1.9	1.8	1.5	1.4	1.4
		58.2	26.9	12.9	7.3	4.5	2.8	2.5	2.0	1.6	1.4	1.3	1.2	1.2
		55.1	24.2	11.5	6.7	3.8	2.5	2.1	1.7	1.6	1.3	1.3	1.2	1.2
		53.4	22.1	10.8	6.3	3.2	2.3	1.8	1.5	1.4	1.3	1.3	1.2	1.2
τ^2	1.0^2	50.2	24.1	13.9	9.5	6.7	4.0	3.1	2.0	1.8	1.6	1.4	1.4	1.4
		37.1	15.8	9.2	6.0	4.2	2.7	2.3	1.8	1.6	1.4	1.3	1.2	1.2
		33.2	14.9	7.9	5.6	3.7	2.4	2.0	1.6	1.5	1.3	1.3	1.2	1.2
		32.9	13.7	6.7	5.4	3.0	2.1	1.8	1.5	1.4	1.3	1.3	1.2	1.2
	1.2^2	30.5	14.2	10.5	7.8	5.8	3.8	3.0	2.0	1.8	1.6	1.4	1.4	1.3
		25.6	10.1	7.2	5.1	3.5	2.6	2.1	1.8	1.6	1.4	1.3	1.2	1.2
		21.9	9.2	6.4	4.8	3.2	2.3	1.9	1.6	1.5	1.3	1.3	1.2	1.2
		20.5	8.5	5.8	4.4	2.8	2.0	1.7	1.5	1.4	1.3	1.3	1.2	1.1
	1.4^2	22.4	13.0	8.5	6.4	5.0	3.5	2.9	1.9	1.7	1.5	1.4	1.3	1.3
		14.0	8.8	5.6	4.3	3.3	2.5	2.1	1.6	1.6	1.4	1.2	1.2	1.2
		12.2	7.2	5.2	4.2	3.0	2.3	1.8	1.6	1.5	1.3	1.2	1.2	1.2
		11.6	6.9	4.8	3.8	2.6	1.8	1.7	1.5	1.4	1.3	1.2	1.2	1.1
	1.6^2	12.9	11.1	7.7	5.5	4.6	3.4	2.8	1.7	1.7	1.5	1.4	1.3	1.3
		9.8	5.5	4.7	3.8	3.2	2.4	1.9	1.6	1.5	1.3	1.2	1.2	1.2
		9.6	5.3	4.4	3.5	2.9	2.1	1.7	1.8	1.4	1.3	1.2	1.2	1.2
		9.1	4.9	4.0	3.3	2.5	1.7	1.6	1.5	1.4	1.3	1.2	1.2	1.1
	1.8^2	8.9	6.9	5.4	4.5	3.8	3.0	2.5	1.6	1.5	1.4	1.4	1.3	1.3
		6.6	4.9	3.7	3.2	2.7	2.3	2.1	1.9	1.5	1.3	1.3	1.3	1.2
		6.1	4.5	3.4	3.0	2.5	2.0	2.0	1.7	1.5	1.3	1.2	1.2	1.2
		5.9	3.9	3.2	2.7	2.4	1.7	1.6	1.5	1.4	1.3	1.2	1.2	1.1
	2.0^2	7.6	5.0	4.2	3.8	3.5	2.9	2.4	1.5	1.4	1.4	1.3	1.3	1.3
		5.5	4.1	3.0	2.7	2.6	1.9	1.8	1.8	1.5	1.3	1.2	1.2	1.2
		5.1	3.8	2.9	2.5	2.4	1.8	1.8	1.6	1.5	1.2	1.2	1.2	1.1
		4.8	3.5	2.8	2.4	2.1	1.5	1.6	1.5	1.4	1.2	1.2	1.2	1.1

Table 5.5: ARLs with unknown in-control parameters case for shifts in $\boldsymbol{\theta}$ with $\boldsymbol{\theta}_j = (\boldsymbol{\beta}_j^T, \delta\sigma)^T$ and $\tau_j^2 = \tau^2/\delta^2$ for $j \geq 1$

	$m=5$ $m=25$ $m=125$ $m=\infty$	δ 1.1^0	1.1^1	1.1^2	1.1^3	1.1^4	1.1^5	1.1^6	1.1^7	1.1^8	1.1^9	1.1^{10}	1.1^{11}	1.1^{12}
τ^2	2.2^2	5.4	4.3	3.6	3.3	3.1	2.6	2.4	2.1	1.9	1.8	1.6	1.5	1.3
		4.1	3.6	2.7	2.3	2.2	2.0	1.8	1.7	1.6	1.5	1.4	1.2	1.2
		3.8	3.3	2.5	2.3	2.2	1.8	1.7	1.6	1.5	1.4	1.3	1.2	1.1
		3.6	3.1	2.4	2.1	1.8	1.7	1.6	1.5	1.3	1.2	1.2	1.2	1.1
	2.4^2	4.5	3.8	3.1	2.8	2.6	2.4	2.2	1.8	1.8	1.7	1.5	1.4	1.3
		3.5	3.0	2.2	2.1	2.0	1.7	1.7	1.6	1.5	1.4	1.3	1.2	1.2
		3.2	2.9	2.2	2.0	1.9	1.7	1.7	1.5	1.5	1.4	1.3	1.2	1.1
		3.0	2.2	2.0	1.9	1.6	1.6	1.6	1.5	1.3	1.2	1.2	1.2	1.1
	2.6^2	4.9	3.2	2.9	2.6	2.4	2.3	2.0	1.7	1.7	1.6	1.4	1.3	1.2
		3.2	2.7	2.1	1.9	1.8	1.7	1.6	1.5	1.4	1.3	1.3	1.2	1.2
		3.0	2.3	1.9	1.8	1.6	1.6	1.5	1.5	1.3	1.3	1.2	1.2	1.1
		2.6	2.0	1.8	1.8	1.5	1.5	1.4	1.4	1.2	1.2	1.2	1.1	1.1
	2.8^2	3.5	3.1	2.4	2.2	2.0	2.0	1.9	1.6	1.6	1.5	1.4	1.3	1.2
		2.8	2.1	1.9	1.7	1.6	1.6	1.5	1.4	1.4	1.3	1.2	1.2	1.2
		2.4	1.9	1.7	1.6	1.6	1.5	1.5	1.4	1.3	1.3	1.2	1.1	1.1
		2.2	1.6	1.5	1.5	1.5	1.4	1.4	1.3	1.3	1.2	1.2	1.1	1.1
	3.0^2	3.2	2.9	2.2	2.0	1.9	1.9	1.8	1.6	1.5	1.4	1.3	1.3	1.3
		2.5	1.9	1.7	1.6	1.5	1.4	1.4	1.4	1.3	1.3	1.2	1.2	1.1
		2.1	1.8	1.6	1.4	1.4	1.4	1.4	1.3	1.3	1.2	1.2	1.1	1.1
		1.8	1.6	1.4	1.4	1.4	1.4	1.3	1.3	1.2	1.2	1.2	1.1	1.1
	3.2^2	2.9	2.7	2.1	1.9	1.8	1.8	1.7	1.5	1.5	1.4	1.3	1.3	1.3
		2.1	1.9	1.6	1.5	1.5	1.4	1.4	1.3	1.3	1.2	1.2	1.2	1.1
		1.9	1.7	1.5	1.4	1.4	1.4	1.3	1.3	1.2	1.2	1.1	1.1	1.1
		1.7	1.5	1.4	1.4	1.4	1.3	1.3	1.2	1.2	1.2	1.1	1.1	1.1
	3.4^2	2.7	2.1	1.9	1.8	1.7	1.7	1.6	1.5	1.5	1.4	1.4	1.3	1.2
		1.8	1.6	1.4	1.3	1.4	1.3	1.3	1.3	1.2	1.2	1.2	1.1	1.1
		1.7	1.5	1.3	1.3	1.3	1.3	1.2	1.2	1.2	1.2	1.1	1.1	1.1
		1.5	1.4	1.3	1.3	1.3	1.2	1.2	1.2	1.2	1.1	1.1	1.1	1.1
	3.6^2	2.0	1.7	1.7	1.7	1.6	1.6	1.5	1.4	1.4	1.3	1.3	1.3	1.2
		1.6	1.5	1.4	1.4	1.3	1.3	1.3	1.2	1.2	1.2	1.2	1.1	1.1
		1.5	1.4	1.3	1.3	1.3	1.3	1.2	1.2	1.2	1.2	1.1	1.1	1.1
		1.4	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.1	1.1	1.1	1.1	1.1
	3.8^2	1.7	1.6	1.5	1.5	1.5	1.5	1.4	1.4	1.3	1.3	1.2	1.3	1.2
		1.5	1.3	1.3	1.2	1.3	1.2	1.2	1.2	1.2	1.2	1.1	1.1	1.1
		1.4	1.3	1.2	1.2	1.2	1.2	1.2	1.2	1.1	1.1	1.1	1.1	1.1
		1.3	1.2	1.2	1.2	1.2	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
	4.0^2	1.5	1.4	1.4	1.4	1.3	1.3	1.3	1.3	1.3	1.2	1.2	1.2	1.2
		1.3	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.1	1.1	1.1	1.1	1.1
		1.2	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
		1.2	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1

Table 5.6: Out of control ARL comparisons between modified (1) EWMA with known in-control parameters case and HEWMA charts for shifts in β_0, where $\boldsymbol{\theta}_j = (\beta_0 + \delta_0\sigma, \beta_1, \sigma)^T$ and $\tau_j^2 = 4\delta_0^2$ for $j \geq 1$.

		ARL_1				
δ_0	τ_j^2	a=1	a=0.75	a=0.5	a=0.25	HEWMA
0.1000	0.0400	174.9	173.2	171.8	170.2	171.1
0.2000	0.1600	118.9	117.3	115.8	112.1	113.1
0.3000	0.3600	71.4	69.7	68.5	67.1	67.2
0.4000	0.6400	40.4	39.5	38.9	38.1	38.1
0.5000	1.0000	23.7	23.0	22.5	21.9	22.1
0.6000	1.4400	14.4	14.1	13.7	13.4	13.8
0.8000	2.5600	6.5	6.4	6.2	6.1	6.1
1.0000	4.0000	3.6	3.6	3.5	3.4	3.4
1.5000	9.0000	1.6	1.5	1.5	1.5	1.5
2.0000	16.000	1.1	1.1	1.1	1.1	1.1

Table 5.7: Out of control ARL comparisons between modified (1) EWMA with known in-control parameters case and HEWMA charts for shifts in β_1, where $\boldsymbol{\theta}_j = (\beta_0, \beta_1 + \delta_1\sigma, \sigma)^T$ and $\tau_j^2 = 120\delta_1^2$ for $j \geq 1$.

		ARL_1				
δ_1	τ_j^2	a=1	a=0.75	a=0.5	a=0.25	HEWMA
0.0250	0.0750	153.6	152.4	152.3	152.2	153.0
0.0375	0.1688	115.6	113.8	112.6	112.3	112.3
0.0500	0.3000	81.6	79.9	78.7	77.6	77.4
0.0625	0.4688	55.8	54.5	53.8	52.8	53.2
0.0750	0.6750	37.7	37.3	36.4	35.3	35.4
0.1000	1.2000	18.6	18.0	17.6	17.2	17.4
0.1250	1.8750	10.0	9.7	9.5	9.3	9.5
0.1500	2.7000	6.1	5.9	5.8	5.6	5.8
0.2000	4.8000	2.9	2.8	2.8	2.7	2.8
0.2500	7.5000	1.8	1.8	1.7	1.7	1.8

Table 5.8: Out of control ARL comparisons between modified (1) EWMA with known in-control parameters case and HEWMA charts for shifts in σ, where $\boldsymbol{\theta}_j = (\beta_0, \beta_1, \delta\sigma)^T$ for $j \geq 1$.

	ARL_1				
δ	a=1	a=0.75	a=0.5	a=0.25	HEWMA
1.10	138.6	148.1	162.6	179.1	51.1
1.15	112.4	126.0	144.0	167.6	30.2
1.20	91.0	105.5	126.9	157.4	20.0
1.25	73.0	87.6	110.9	145.9	14.6
1.30	59.2	72.7	96.3	135.9	11.0
1.40	38.6	49.7	71.2	114.7	7.0
1.60	18.2	24.0	37.8	77.7	4.0
1.80	10.4	13.3	20.7	49.8	2.8
2.20	4.9	5.9	8.5	20.2	1.8
2.60	3.2	3.6	4.8	9.8	1.4
3.00	2.4	2.7	3.3	5.9	1.3

Table 5.9: Out of control ARL comparisons between modified (2) MEWMA with known in-control parameters case, HEWMA, ZTW and KMW charts for shifts in β_0, where $\boldsymbol{\theta}_j = (\beta_0 + \delta_0\sigma, \beta_1, \sigma)^T$.

	ARL_1					
	unconstraint	unconstraint	unconstraint	HEWMA	ZTW	KMW
δ_0	a=1	a=0.75	a=0.5			
0.1000	131.8	127.3	126.2	171.1	131.5	131.5
0.2000	59.7	58.0	57.1	113.1	59.9	59.1
0.3000	28.5	27.7	27.3	67.2	29.6	28.3
0.4000	16.5	16.0	15.7	38.1	17.2	16.2
0.5000	10.2	9.8	9.6	22.1	11.5	10.7
0.6000	7.5	7.3	7.1	13.8	8.5	7.9
0.8000	4.0	3.8	3.7	6.1	5.8	5.1
1.0000	2.8	2.7	2.6	3.4	4.1	3.8
1.5000	1.5	1.5	1.4	1.4	2.6	2.4
2.0000	1.1	1.1	1.1	1.1	2.0	1.9

Table 5.10: Out of control ARL comparisons between modified (2) MEWMA with known in-control parameters case, HEWMA, ZTW and KMW charts for shifts in β_1, where $\boldsymbol{\theta}_j = (\beta_0, \beta_1 + \delta_1\sigma, \sigma)^T$.

	ARL$_1$					
	unconstraint	unconstraint	unconstraint	HEWMA	ZTW	KMW
δ_1	a=1	a=0.75	a=0.5			
0.0250	99.9	98.9	98.2	153.0	99.0	101.6
0.0375	57.3	56.3	55.9	112.3	57.4	61.0
0.0500	34.8	33.9	32.1	77.4	35.0	36.5
0.0625	22.9	22.0	21.5	53.2	23.1	24.6
0.0750	16.0	15.4	15.0	35.4	16.4	17.0
0.1000	8.5	8.2	8.0	17.4	9.8	10.3
0.1250	5.8	5.6	5.2	9.5	6.9	7.2
0.1500	4.5	4.5	4.3	5.8	5.3	5.5
0.2000	2.7	2.6	2.5	2.8	3.7	3.8
0.2500	1.7	1.7	1.6	1.8	2.9	2.9

Table 5.11: Out of control ARL comparisons between modified (2) MEWMA with known in-control parameters case, HEWMA, ZTW and KMW charts for shifts in σ, where $\boldsymbol{\theta}_j = (\beta_0, \beta_1, \delta\sigma)^T$ for $j \geq 1$.

	ARL$_1$						
	unconstraint	unconstraint	unconstraint	constraint	HEWMA	ZTW	KMW
δ	a=1	a=0.75	a=0.5	a=1			
1.10	53.1	58.1	66.1	50.0	51.1	76.2	72.8
1.15	32.2	38.2	45.0	30.1	30.2	48.7	48.1
1.20	20.8	24.3	29.4	19.8	20.0	33.2	33.5
1.25	15.2	17.8	21.8	14.0	14.6	24.1	24.9
1.30	11.0	12.6	15.4	10.0	11.0	18.4	19.4
1.40	6.9	7.8	9.3	6.4	7.0	12.1	12.7
1.60	3.9	4.3	4.8	3.8	4.0	7.0	7.2
1.80	2.8	3.0	3.2	2.4	2.8	4.9	5.1
2.20	1.7	1.8	2.0	1.6	1.8	3.1	3.2
2.60	1.4	1.5	1.7	1.3	1.4	2.3	2.5
3.00	1.2	1.3	1.4	1.2	1.3	1.9	2.1

Appendix

A.1

$$\ell_j(\boldsymbol{\theta}) = \ell_j(\hat{\boldsymbol{\theta}}_j) + \left[\ell_j'(\hat{\boldsymbol{\theta}}_j)\right]^T (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_j) + \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_j)^T \ell_j'' \left(\hat{\boldsymbol{\theta}}_j + \hat{\eta}_j(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_j)\right) (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_j), \quad (A.1)$$

where $0 < \hat{\eta}_j < 1$ and

$$\ell_j'' \left(\hat{\boldsymbol{\theta}}_j + \hat{\eta}_j(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_j)\right) = \ell_j''(\boldsymbol{\theta}) + O_p(\sqrt{n_j}) = -\mathrm{Cov}_{\boldsymbol{\theta}}^{-1}(\ell_j'(\boldsymbol{\theta})) + O_p(\sqrt{n_j}). \quad (A.2)$$

$$\ell_j'(\boldsymbol{\theta}) = \ell_j'(\hat{\boldsymbol{\theta}}_j) + \ell_j''(\hat{\boldsymbol{\theta}}_j)(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_j) + O_p(1) = \mathrm{Cov}_{\boldsymbol{\theta}}(\ell_j'(\boldsymbol{\theta}))(\hat{\boldsymbol{\theta}}_j - \boldsymbol{\theta}) + O_p(1),$$

$$\hat{\boldsymbol{\theta}}_j - \boldsymbol{\theta} = \mathrm{Cov}_{\boldsymbol{\theta}}^{-1}\left(\ell_j'(\boldsymbol{\theta})\right) \ell_j'(\boldsymbol{\theta}_j) + O_p\left(\frac{1}{n_j}\right). \quad (A.3)$$

By (A.1), (A.2) and (A.3)

$$\begin{aligned}
& 2\left[\ell_j'(\hat{\boldsymbol{\theta}}_j) - \ell_j'(\boldsymbol{\theta})\right] \\
= \; & (\hat{\boldsymbol{\theta}}_j - \boldsymbol{\theta})^T \mathrm{Cov}_{\boldsymbol{\theta}}(\ell_j'(\boldsymbol{\theta}))(\hat{\boldsymbol{\theta}}_j - \boldsymbol{\theta}) + O_p\left(\frac{1}{\sqrt{n_j}}\right) \\
= \; & \left[\ell_j'(\boldsymbol{\theta})\right]^T \mathrm{Cov}_{\boldsymbol{\theta}}^{-1}\left(\ell_j'(\boldsymbol{\theta})\right) \mathrm{Cov}_{\boldsymbol{\theta}}\left(\ell_j'(\boldsymbol{\theta})\right) \mathrm{Cov}_{\boldsymbol{\theta}}^{-1}\left(\ell_j'(\boldsymbol{\theta})\right) \ell_j'(\boldsymbol{\theta}_j) + O_p\left(\frac{1}{\sqrt{n_j}}\right) \\
= \; & W_j + O_p\left(\frac{1}{\sqrt{n_j}}\right).
\end{aligned}$$

A.2

When the process is in control,

$$H_{j1} \sim \chi^2_p$$

and

$$H_{j2} \sim \chi^2_{n_j-p}.$$

$$\mathrm{E}_{\theta}(W_j) = \mathrm{E}_{\boldsymbol{\theta}}(H_{j1} + \frac{1}{2n_j}(H_{j1} + H_{j2} - n_j)^2) = p + \frac{\mathrm{Var}(H_{j1} + H_{j2})}{2n_j} = p + 1.$$

$$\begin{aligned}
&\mathrm{Var}_{\boldsymbol{\theta}}(W_j) \\
= \;& \frac{1}{4n_j^2}\mathrm{E}_{\boldsymbol{\theta}}\left((H_{j1}^2 + H_{j2}^2 + 2H_{j1}H_{j2} - 2n_jH_{j2} + n_j^2)^2\right) - (p+1)^2 \\
= \;& \frac{1}{4n_j^2}\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^4 + H_{j2}^4 + 6H_{j1}^2H_{j2}^2 + 4H_{j1}^3H_{j2} + 4H_{j1}H_{j2}^3 - 4n_jH_{j1}^2H_{j2} - 8n_jH_{j1}H_{j2}^2 \\
& -4n_jH_{j2}^3 + 2n_j^2H_{j1}^2 + 6n_j^2H_{j2}^2 + 4n_j^2H_{j1}H_{j2} - 4n_j^3H_{j2} + n_j^4) - (p+1)^2 \\
= \;& \frac{1}{4n_j^2}[\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^4) + \mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^4) + 6\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^2)\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^2) + 4\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^3)\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}) \\
& +4\mathrm{E}_{\boldsymbol{\theta}}(H_{j1})\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^3) - 4n_j\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^2)\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}) - 8n_j\mathrm{E}_{\boldsymbol{\theta}}(H_{j1})\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^2) - 4n_j\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^3) \\
& +2n_j^2\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^2) + 6n_j^2\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^2) + 4n_j^2\mathrm{E}_{\boldsymbol{\theta}}(H_{j1})\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}) - 4n_j^3\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}) + n_j^4\,] - (p+1)^2 \\
= \;& 2p + \frac{8p}{n_j} + \frac{12}{n_j} + 2
\end{aligned}$$

where

$$
\begin{aligned}
&\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}) = p,\\
&\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^2) = p(p+2),\\
&\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^3) = p(p+2)(p+4),\\
&\mathrm{E}_{\boldsymbol{\theta}}(H_{j1}^4) = p(p+2)(p+4)(p+6),\\
&\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}) = n_j - p,\\
&\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^2) = (n_j - p)(n_j - p + 2),\\
&\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^3) = (n_j - p)(n_j - p + 2)(n_j - p + 4),\\
&\mathrm{E}_{\boldsymbol{\theta}}(H_{j2}^4) = (n_j - p)(n_j - p + 2)(n_j - p + 4)(n_j - p + 6)
\end{aligned}
$$

B.1

$$\mathrm{E}_{\boldsymbol{\theta}}(W_j^*) = \mathrm{E}_{\boldsymbol{\theta}}(W_j) - \mathrm{E}_{\boldsymbol{\theta}_0}\left(W_j(\boldsymbol{\theta}_j^*)\right),$$

where $\mathrm{E}_{\boldsymbol{\theta}}(W_j) = p + 1$ from Appendix A.2 and

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{\theta}_0}\left(W_j(\boldsymbol{\theta}_j^*)\right) \\
&= \frac{n_j}{2}\mathrm{E}_{\boldsymbol{\theta}}\left(\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) \\
&= \frac{n_j}{2}\left\{E_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}^2}{n_j^2}, 1\right)\right) - 2E_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}}{n_j}, 1\right)\right) + 1\right\}
\end{aligned}$$

with

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}^2}{n_j^2}, 1\right)\right) \\
&= \int_0^{n_j}\left(\frac{x}{n_j}\right)^2 f_{\chi^2_{n_j-p}}(x)dx + \int_{n_j}^{\infty} 1 \cdot f_{\chi^2_{n_j-p}}(x)dx \\
&= \int_0^{n_j}\left(\frac{x}{n_j}\right)^2 \cdot \frac{x^{(n_j-p)/2-1}\cdot e^{-x/2}}{\Gamma((n_j-p)/2)\cdot 2^{(n_j-p)/2}}dx + \left[1 - \int_0^{n_j} f_{\chi^2_{n_j-p}}(x)dx\right] \\
&= \frac{2^2\Gamma((n_j-p)/2+2)}{n_j^2\Gamma((n_j-p)/2)}\cdot\int_0^{n_j}\frac{x^{(n_j-p)/2+2-1}\cdot e^{-x/2}}{\Gamma((n_j-p)/2+2)\cdot 2^{(n_j-p)/2+2}}dx + \left[1 - F_{\chi^2_{n_j-p}}(n_j)\right] \\
&= \frac{1}{n_j^2}(n_j-p)(n_j-p+2)\cdot F_{\chi^2_{n_j-p+4}}(n_j) + \left[1 - F_{\chi^2_{n_j-p}}(n_j)\right]
\end{aligned}$$

and

$$
\begin{aligned}
&\mathrm{E}_{\boldsymbol{\theta}_0}\left(\min\left(\frac{H_{j2}}{n_j},1\right)\right)\\
&=\int_0^{n_j}\left(\frac{x}{n_j}\right)f_{\chi^2_{n_j-p}}(x)dx+\int_{n_j}^{\infty}1\cdot f_{\chi^2_{n_j-p}}(x)dx\\
&=\int_0^{n_j}\left(\frac{x}{n_j}\right)\cdot\frac{x^{(n_j-p)/2-1}\cdot e^{-x/2}}{\Gamma((n_j-p)/2)\cdot 2^{(n_j-p)/2}}dx+\left[1-\int_0^{n_j}f_{\chi^2_{n_j-p}}(x)dx\right]\\
&=\frac{2\Gamma((n_j-p)/2+1)}{n_j\Gamma((n_j-p)/2)}\cdot\int_0^{n_j}\frac{x^{(n_j-p)/2+1-1}\cdot e^{-x/2}}{\Gamma((n_j-p)/2+1)\cdot 2^{(n_j-p)/2+1}}dx+\left[1-F_{\chi^2_{n_j-p}}(n_j)\right]\\
&=\frac{1}{n_j}(n_j-p)\cdot F_{\chi^2_{n_j-p+2}}(n_j)+\left[1-F_{\chi^2_{n_j-p}}(n_j)\right].
\end{aligned}
$$

$$
\mathrm{Var}_{\boldsymbol{\theta}}(W_j^*)=\mathrm{Var}_{\boldsymbol{\theta}}(W_j-W_j(\boldsymbol{\theta}_j^*))=\mathrm{Var}_{\boldsymbol{\theta}}(W_j)+\mathrm{Var}_{\boldsymbol{\theta}}(W_j(\boldsymbol{\theta}_j^*))-2\mathrm{Cov}_{\boldsymbol{\theta}}(W_j,W_j(\boldsymbol{\theta}_j^*)),
$$

where $\mathrm{Var}_{\boldsymbol{\theta}}(W_j)$ is given in Appendix A.2 and

$$
\begin{aligned}
\mathrm{Var}_{\boldsymbol{\theta}_0}(W_j(\boldsymbol{\theta}_j^*)) &= \frac{n_j^2}{4}\mathrm{Var}_{\boldsymbol{\theta}_0}\left(\left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right)\\
&= \frac{n_j^2}{4}\mathrm{E}_{\boldsymbol{\theta}_0}\left(\left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^4\right)-\frac{n_j^2}{4}\left[\mathrm{E}_{\boldsymbol{\theta}_0}\left(\left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right)\right]^2,
\end{aligned}
$$

with

$$
\begin{aligned}
&\mathrm{E}_{\boldsymbol{\theta}}\left(\left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^4\right)\\
= \;&\mathrm{E}_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}^4}{n_j^4},1\right)\right)-4E_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}^3}{n_j^3},1\right)\right)+6E_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}^2}{n_j^2},1\right)\right)\\
&-4\mathrm{E}_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}}{n_j},1\right)\right)+1.
\end{aligned}
$$

Here

$$
\begin{aligned}
& \mathrm{E}_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}^4}{n_j^4},1\right)\right) \\
= \; & \frac{1}{n_j^4}(n_j-p)(n_j-p+2)(n_j-p+4)(n_j-p+6)\cdot F_{\chi^2_{n_j-p+8}}(n_j)+\left[1-F_{\chi^2_{n_j-p}}(n_j)\right] \\
& \mathrm{E}_{\boldsymbol{\theta}}\left(\min\left(\frac{H_{j2}^3}{n_j^3},1\right)\right) \\
= \; & \frac{1}{n_j^3}(n_j-p)(n_j-p+2)(n_j-p+4)\cdot F_{\chi^2_{n_j-p+6}}(n_j)+\left[1-F_{\chi^2_{n_j-p}}(n_j)\right].
\end{aligned}
$$

Finally,

$$
\begin{aligned}
& \mathrm{Cov}_{\boldsymbol{\theta}}(W_j, W_j(\boldsymbol{\theta}_j^*)) \\
= \; & \frac{1}{4}\mathrm{Cov}_{\boldsymbol{\theta}}\left((H_{j1}+H_{j2}-n_j)^2, \left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right) \\
= \; & \frac{1}{4}\mathrm{Cov}_{\boldsymbol{\theta}}\left(2H_{j1}H_{j2}+H_{j2}^2-2n_jH_{j2}, \left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right) \\
= \; & \frac{1}{2}\mathrm{Cov}_{\boldsymbol{\theta}}\left(H_{j1}H_{j2}, \left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right)+\frac{1}{4}\mathrm{Cov}_{\boldsymbol{\theta}}\left(H_{j2}^2, \left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right) \\
& -\frac{n_j}{2}\mathrm{Cov}_{\boldsymbol{\theta}}\left(H_{j2}, \left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right),
\end{aligned}
$$

where

$$
\begin{aligned}
& \mathrm{Cov}_{\boldsymbol{\theta}}\left(H_{j1}H_{j2}, \left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) \\
= \; & \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j1}H_{j2}\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) - \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j1}H_{j2}\right) \cdot \mathrm{E}_{\boldsymbol{\theta}}\left(\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) \\
= \; & p \cdot \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) - p(n_j - p) \cdot \mathrm{E}_{\boldsymbol{\theta}}\left(\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right),
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{Cov}_{\boldsymbol{\theta}}\left(H_{j2}^2, \left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) \\
= \; & \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}^2\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) - \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}^2\right) \cdot \mathrm{E}_{\boldsymbol{\theta}}\left(\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) \\
= \; & \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}^2\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) - \left[2(n_j - p) + (n_j - p)^2\right] \cdot \mathrm{E}_{\boldsymbol{\theta}}\left(\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right),
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{Cov}_{\boldsymbol{\theta}}\left(H_{j2}, \left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) \\
= \; & \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) - \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}\right) \cdot \mathrm{E}_{\boldsymbol{\theta}}\left(\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) \\
= \; & \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right) - (n_j - p) \cdot \mathrm{E}_{\boldsymbol{\theta}_0}\left(\left[\min\left(\frac{H_{j2}}{n_j}, 1\right) - 1\right]^2\right).
\end{aligned}
$$

Here

$$
\begin{aligned}
& \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}\left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right) \\
= & \int_0^{n_j} x\left(\frac{x}{n_j}-1\right)^2 f_{\chi^2_{n_j-p}}(x)dx \\
= & \frac{1}{n_j}\int_0^{n_j} x^3 f_{\chi^2_{n_j-p}}(x)dx-\frac{2}{n_j}\int_0^{n_j} x^2 f_{\chi^2_{n_j-p}}(x)dx+\int_0^{n_j} x f_{\chi^2_{n_j-p}}(x)dx
\end{aligned}
$$

$$
\begin{aligned}
& \mathrm{E}_{\boldsymbol{\theta}}\left(H_{j2}^2\left[\min\left(\frac{H_{j2}}{n_j},1\right)-1\right]^2\right) \\
= & \int_0^{n_j} x^2\left(\frac{x}{n_j}-1\right)^2 f_{\chi^2_{n_j-p}}(x)dx \\
= & \frac{1}{n_j^2}\int_0^{n_j} x^4 f_{\chi^2_{n_j-p}}(x)dx-\frac{2}{n_j}\int_0^{n_j} x^3 f_{\chi^2_{n_j-p}}(x)dx+\int_0^{n_j} x^2 f_{\chi^2_{n_j-p}}(x)dx
\end{aligned}
$$

with

$$
\begin{aligned}
& \int_0^{n_j} x^k f_{\chi^2_{n_j-p}}(x)dx \\
= & (n_j-p)(n_j-p+2)\cdots[n_j-p+2(k-1)]\cdot F_{\chi^2_{n_j-p+2k}}(n_j),\ \ k=1,2,\ldots.
\end{aligned}
$$

C.1

Rewrite $W_{0,j}$ in section 2.4 as

$$
\begin{aligned}
W_{0,j} &= \frac{(\boldsymbol{y}_0 - X_0\tilde{\boldsymbol{\beta}})^T X_0 (X_0^T X_0)^{-1} X_0^T (\boldsymbol{y}_0 - X_0\tilde{\boldsymbol{\beta}})}{\tilde{\sigma}^2} \\
&+ \frac{(\boldsymbol{y}_j - X_j\tilde{\boldsymbol{\beta}}_j)^T X_j (X_j^T X_j)^{-1} X_j^T (\boldsymbol{y}_j - X_j\tilde{\boldsymbol{\beta}}_j)}{\tilde{\sigma}_j^2} \\
&+ \frac{1}{2n_0}\left[\frac{\|\boldsymbol{y}_0 - X_0\tilde{\boldsymbol{\beta}}\|^2 - n_0\tilde{\sigma}^2}{\tilde{\sigma}^2}\right]^2 + \frac{1}{2n_j}\left[\frac{\|\boldsymbol{y}_j - X_j\tilde{\boldsymbol{\beta}}_j\|^2 - n_j\tilde{\sigma}_j^2}{\tilde{\sigma}_j^2}\right]^2
\end{aligned}
$$

note that

$$
\begin{aligned}
\varepsilon_{0j}^{(1)} &\equiv \boldsymbol{y}_0 - X_0\tilde{\boldsymbol{\beta}} = \varepsilon_0 - X_0(X_0^T X_0 + X_j^T X_j)^{-1}(X_0^T\varepsilon_0 + X_j^T\varepsilon_j), \\
\varepsilon_{0j}^{(2)} &\equiv \boldsymbol{y}_0 - X_j\tilde{\boldsymbol{\beta}}_j = \varepsilon_j - X_j(X_0^T X_0 + X_j^T X_j)^{-1}(X_0^T\varepsilon_0 + X_j^T\varepsilon_j), \\
\tilde{\sigma}^2 &= \tilde{\sigma}_j^2 = \frac{\|\ \boldsymbol{y}_0 - X_0\tilde{\boldsymbol{\beta}}\ \|^2 + \|\ \boldsymbol{y}_j - X_j\tilde{\boldsymbol{\beta}}_j\ \|^2}{n_0 + n_j} = \frac{\|\ \varepsilon_{0j}^{(1)}\ \|^2 + \|\ \varepsilon_{0j}^{(2)}\ \|^2}{n_0 + n_j}
\end{aligned}
$$

then

$$
\begin{aligned}
W_{0,j} &= \frac{(\varepsilon_{0j}^{(1)})^T X_0 (X_0^T X_0)^{-1} X_0^T (\varepsilon_{0j}^{(1)}) + (\varepsilon_{0j}^{(2)})^T X_j (X_j^T X_j)^{-1} X_j^T (\varepsilon_{0j}^{(2)})}{\tilde{\sigma}_0^2} \\
&+ \frac{1}{2n_0}\left[\frac{\|\varepsilon_{0j}^{(1)}\|^2 - n_0\tilde{\sigma}^2}{\tilde{\sigma}^2}\right]^2 + \frac{1}{2n_j}\left[\frac{\|\varepsilon_{0j}^{(2)}\|^2 - n_j\tilde{\sigma}_j^2}{\tilde{\sigma}_j^2}\right]^2 \\
&= \frac{(n_0 + n_j)\left[(\varepsilon_{0j}^{(1)})^T X_0 (X_0^T X_0)^{-1} X_0^T (\varepsilon_{0j}^{(1)}) + (\varepsilon_{0j}^{(2)})^T X_j (X_j^T X_j)^{-1} X_j^T (\varepsilon_{0j}^{(2)})\right]}{\|\varepsilon_{0j}^{(1)}\|^2 + \|\varepsilon_{0j}^{(2)}\|^2} \\
&+ \frac{(n_0 + n_j)^3}{2n_0 n_j\left(\|\varepsilon_{0j}^{(1)}\|^2 + \|\varepsilon_{0j}^{(2)}\|^2\right)^2}\left[\|\varepsilon_{0j}^{(2)}\|^2 - \frac{n_j\left(\|\varepsilon_{0j}^{(1)}\|^2 + \|\varepsilon_{0j}^{(2)}\|^2\right)}{n_0 + n_j}\right]^2
\end{aligned}
$$

Let

$$
\begin{aligned}
Z_0 &\equiv (X_0^T X_0)^{-1/2} X_0^T \varepsilon_0, \\
Z_j &\equiv (X_j^T X_j)^{-1/2} X_j^T \varepsilon_j, \\
\text{and} \quad & Z_0, Z_j \overset{iid}{\sim} N_p(0_{p\times 1}, I_p), \\
A &\equiv \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \\
\text{where } a_{11} &= X_0^T X_0 + X_j^T X_j)^{-1/2} (X_0^T X_0)^{1/2} \\
a_{12} &= (X_0^T X_0 + X_j^T X_j)^{-1/2} (X_j^T X_j)^{1/2} \\
a_{21} &= \left[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}\right]^{-1/2} (X_0^T X_0)^{-1/2} \\
a_{22} &= -\left[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}\right]^{-1/2} (X_j^T X_j)^{-1/2} \\
\text{and} \quad & A^{-1} = A^T, \\
\begin{pmatrix} Z_{0j}^{(1)} \\ Z_{0j}^{(2)} \end{pmatrix} &\equiv A \begin{pmatrix} Z_0 \\ Z_j \end{pmatrix} \\
\text{and} \quad & Z_{0j}^{(1)},\ Z_{0j}^{(2)} \overset{iid}{\sim} N_p(0_{p\times 1}, I_p), \quad \begin{pmatrix} Z_0 \\ Z_j \end{pmatrix} = A^T \begin{pmatrix} Z_{0j}^{(1)} \\ Z_{0j}^{(2)} \end{pmatrix}, \\
H_0 &\equiv \|\varepsilon_0\|^2 - \|Z_0\|^2 \sim \chi^2_{n_0 - p} \\
H_j &\equiv \|\varepsilon_j\|^2 - \|Z_j\|^2 \sim \chi^2_{n_j - p} \\
\text{and} \quad & H_0,\ H_j,\ Z_{0j}^{(1)},\ Z_{0j}^{(2)} \text{ are independent.}
\end{aligned}
$$

Then

$$\begin{aligned}
\varepsilon_{0j}^{(1)} &= \varepsilon_0 - X_0(X_0^T X_0 + X_j^T X_j)^{-1}(X_0^T\varepsilon_0 + X_j^T\varepsilon_j) \\
&= \varepsilon_0 - X_0(X_0^T X_0 + X_j^T X_j)^{-1/2} Z_{0j}^{(1)} \\
\text{then } \|\varepsilon_{0j}^{(1)}\|^2 &= \|\varepsilon_0\|^2 - 2Z_0^T(X_0^T X_0)^{1/2}(X_0^T X_0 + X_j^T X_j)^{-1/2} Z_{0j}^{(1)} \\
&\quad + \|(X_0^T X_0)^{1/2}(X_0^T X_0 + X_j^T X_j)^{-1/2} Z_{0j}^{(1)}\|^2 \\
&= H_0 + \|Z_0 - (X_0^T X_0)^{1/2}(X_0^T X_0 + X_j^T X_j)^{-1/2} Z_{0j}^{(1)}\|^2 \\
&= H_0 + \|(X_0^T X_0)^{-1/2}\left[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}\right]^{-1/2} Z_{0j}^{(2)}\|^2.
\end{aligned}$$

And

$$\begin{aligned}
&\operatorname{cov}\left((X_0^T X_0)^{-1/2}\left[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}\right]^{-1/2} Z_{0j}^{(2)}\right) \\
&= (X_0^T X_0)^{-1/2}\left[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}\right]^{-1}(X_0^T X_0)^{-1/2} \\
&= \left\{(X_0^T X_0)^{1/2}\left[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}\right](X_0^T X_0)^{1/2}\right\}^{-1} \\
&= \left[I_p + (X_0^T X_0)^{1/2}(X_j^T X_j)^{-1}(X_0^T X_0)^{1/2}\right]^{-1} \\
&= \left(I_p + P_{0j}\Lambda_{0j}^{-1}P_{0j}^T\right)^{-1} = \left[P_{0j}(I_p + \Lambda_{0j}^{-1})P_{0j}^T\right]^{-1} = P_{0j}(I_p + \Lambda_{0j}^{-1})^{-1}P_{0j}^T
\end{aligned}$$

where $P_{0j}\Lambda_{0j}^{-1}P_{0j}^T$ is eigendecomposition of $(X_0^T X_0)^{1/2}(X_j^T X_j)^{-1}(X_0^T X_0)^{1/2}$ and

$$P_{0j}^{-1} = P_{0j}^T\ ,\ \Lambda_{0j} \equiv diag\{\lambda_{0j1}, \cdots, \lambda_{0jp}\}, \lambda_{0j1} \geq \cdots \geq \lambda_{0jp} \geq 0$$

and $\lambda_{0j1}, \cdots, \lambda_{0jp}$ are the eigenvalues of $(X_0^T X_0)^{1/2}(X_j^T X_j)^{-1}(X_0^T X_0)^{1/2}$, $P_{0j} \equiv (P_{0j1}, \cdots, P_{0jp})$, P_{0jk} is the eigenvector of the eigenvalue λ_{0jk}, $k = 1, 2, \ldots, p$.

Let

$$
\begin{aligned}
Z_{0j} &\equiv (I_p+\Lambda_{0j}^{-1})^{1/2}P_{0j}^T(X_0^TX_0)^{-1/2}\left[(X_0^TX_0)^{-1}+(X_j^TX_j)^{-1}\right]^{-1/2}Z_{0j}^{(2)} \\
&\sim N_p(O_{p\times 1},I_p), \|Z_{0j}^{(2)}\|^2=\|Z_{0j}\|^2 \\
\text{then } \|\varepsilon_{0j}^{(1)}\|^2 &= H_0+\|\,(I_p+\Lambda_{0j}^{-1})^{-1/2}Z_{0j}\,\|^2=H_0+\sum_{k=1}^{p}\frac{Z_{0jk}^2}{1+1/\lambda_{0jk}} \\
&= H_0+\sum_{k=1}^{p}\frac{\lambda_{0jk}Z_{0jk}^2}{1+\lambda_{0jk}}. \ (C.1)
\end{aligned}
$$

And

$$
\begin{aligned}
\varepsilon_{0j}^{(2)} &\equiv \varepsilon_j-X_j(X_0^TX_0+X_j^TX_j)^{-1}(X_0^T\varepsilon_0+X_j^T\varepsilon_j) \\
&= \varepsilon_j-X_j(X_0^TX_0+X_j^TX_j)^{-1/2}Z_{0j}^{(1)} \\
\text{then } \|\varepsilon_{0j}^{(2)}\|^2 &= \|\varepsilon_j\|^2-2Z_j^T(X_j^TX_j)^{1/2}(X_0^TX_0+X_j^TX_j)^{-1/2}Z_{0j}^{(1)} \\
&\quad +\|(X_j^TX_j)^{1/2}(X_0^TX_0+X_j^TX_j)^{-1/2}Z_{0j}^{(1)}\|^2 \\
&= H_j+\|Z_j-(X_j^TX_j)^{1/2}(X_0^TX_0+X_j^TX_j)^{-1/2}Z_{0j}^{(1)}\|^2 \\
&= H_j+\|(X_j^TX_j)^{-1/2}\left[(X_0^TX_0)^{-1}+(X_j^TX_j)^{-1}\right]^{-1/2}Z_{0j}^{(2)}\|^2 \\
&= H_j+\|Z_{0j}^{(2)}\|^2-\|(X_0^TX_0)^{-1/2}\left[(X_0^TX_0)^{-1}+(X_j^TX_j)^{-1}\right]^{-1/2}Z_{0j}^{(2)}\|^2 \\
&= H_j+\|Z_{0j}\|^2-\sum_{k=1}^{p}\frac{\lambda_{0jk}Z_{0jk}^2}{1+\lambda_{0jk}} \\
&= H_j+\sum_{k=1}^{p}\frac{Z_{0jk}^2}{1+\lambda_{0jk}}. \ (C.2)
\end{aligned}
$$

Hence

$$
\|\varepsilon_{0j}^{(1)}\|^2+\|\varepsilon_{0j}^{(2)}\|^2=H_0+H_j+\sum_{k=1}^{p}Z_{0jk}^2=H_0+H_j+\|Z_{0j}\|^2\sim\chi^2_{n_0+n_j-p}. \ (C.3)
$$

And

$$(\varepsilon_{0j}^{(1)})^T X_0 (X_0^T X_0)^{-1} X_0^T (\varepsilon_{0j}^{(1)}) + (\varepsilon_{0j}^{(2)})^T X_j (X_j^T X_j)^{-1} X_j^T (\varepsilon_{0j}^{(2)})$$

$$\begin{aligned}
&= \|(X_0^T X_0)^{-1/2} X_0^T \varepsilon_{0j}^{(1)}\|^2 + \|(X_j^T X_j)^{-1/2} X_j^T \varepsilon_{0j}^{(2)}\|^2 \\
&= \|Z_0 - (X_0^T X_0)^{1/2} (X_0^T X_0 + X_j^T X_j)^{-1/2} Z_{0j}^{(1)}\|^2 \\
&\quad + \|Z_j - (X_j^T X_j)^{1/2} (X_0^T X_0 + X_j^T X_j)^{-1/2} Z_{0j}^{(1)}\|^2 \\
&= \|(X_0^T X_0)^{-1/2} \left[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}\right]^{-1/2} Z_{0j}^{(2)}\|^2 \\
&\quad + \|(X_j^T X_j)^{-1/2} \left[(X_0^T X_0)^{-1} + (X_j^T X_j)^{-1}\right]^{-1/2} Z_{0j}^{(2)}\|^2 \\
&= \|Z_{0j}^{(2)}\|^2 = \|Z_{0j}\|^2 = \sum_{k=1}^{p} Z_{0jk}^2. \quad (C.4)
\end{aligned}$$

By (C.1), (C.2), (C.3), (C.4)

$$\begin{aligned}
W_{0,j} &= \frac{(n_0+n_j)\|Z_{0j}\|^2}{H_0+H_j+\|Z_{0j}\|^2} \\
&+ \frac{(n_0+n_j)^3}{2n_0 n_j \left(H_0+H_j+\|Z_{0j}\|^2\right)^2} \left[H_j + \sum_{k=1}^{p} \frac{Z_{0jk}^2}{1+\lambda_{0jk}} - \frac{n_j\left(H_0+H_j+\|Z_{0j}\|^2\right)}{n_0+n_j}\right]^2.
\end{aligned}$$

Let

$$\begin{aligned}
D_{0j0} &\equiv \frac{H_0}{H_0+H_j+\|Z_{0j}\|^2}, \\
D_{0jk} &\equiv \frac{Z_{0jk}^2}{H_0+H_j+\|Z_{0j}\|^2}, k = 1, 2, \ldots, p, \\
D_{0j,p+1} &\equiv \frac{H_j}{H_0+H_j+\|Z_{0j}\|^2}, \\
&(D_{0j0}, D_{0j1}, \cdots, D_{0jp}, D_{0j,p+1})^T \sim \text{Dirichlet}\left(\frac{n_0-p}{2}, \frac{1}{2}, \cdots, \frac{1}{2}, \frac{n_j-p}{2}\right).
\end{aligned}$$

And

$$E(D_{0j0}) = \frac{n_0 - p}{n_0 + n_j - p}, \mathrm{Var}(D_{0j0}) = \frac{2n_j(n_0 - p)}{(n_0 + n_j - p)^2(n_0 + n_j - p + 2)},$$
$$E(D_{0jk}) = \frac{1}{n_0 + n_j - p}, \mathrm{Var}(D_{0jk}) = \frac{2(n_0 + n_j - p - 1)}{(n_0 + n_j - p)^2(n_0 + n_j - p + 2)}, k = 1, 2, \ldots, p.$$
$$E(D_{0j,p+1}) = \frac{n_j - p}{n_0 + n_j - p}, \mathrm{Var}(D_{0j,p+1}) = \frac{2n_0(n_j - p)}{(n_0 + n_j - p)^2(n_0 + n_j - p + 2)}.$$

Then

$$W_{0,j} = (n_0 + n_j) \sum_{k=1}^{p} D_{0jk} + \frac{(n_0 + n_j)^3}{2n_0 n_j} \left[D_{0j,p+1} + \sum_{k=1}^{p} \frac{D_{0jk}}{1 + \lambda_{0jk}} - \frac{n_j}{n_0 + n_j} \right]^2.$$
$$E(W_{0,j}) = \frac{(n_0 + n_j)p}{n_0 + n_j - p} + \frac{(n_0 + n_j)^3}{2n_0 n_j} \left\{ (e_1)^2 + v_1 \right\}.$$

Where

$$\begin{aligned} e_1 &= E\left(D_{0j,p+1} + \sum_{k=1}^{p} \frac{D_{0jk}}{1+\lambda_{0jk}} - \frac{n_j}{n_0+n_j}\right) \\ &= \frac{n_j - p}{n_{0+n_j-p}} + \frac{1}{n_0+n_j-p} \sum_{k=1}^{p} \frac{1}{\lambda_{0jk}} - \frac{n_j}{n_0+n_j}, \ (C.5) \end{aligned}$$

and

$$\begin{aligned} v_1 &= \text{Var}\left(D_{0j,p+1} + \sum_{k=1}^{p} \frac{D_{0jk}}{1+\lambda_{0jk}} - \frac{n_j}{n_0+n_j}\right) = \text{Var}\left(D_{0j,p+1} + \sum_{k=1}^{p} \frac{D_{0jk}}{1+\lambda_{0jk}}\right) \\ &= \text{Var}(D_{0j,p+1}) + \sum_{k=1}^{p} \left(\frac{1}{1+\lambda_{0jk}}\right)^2 \text{Var}(D_{0jk}) \\ &\quad +2 \sum_{k=1}^{p} \left(\frac{1}{1+\lambda_{0jk}}\right) \text{Cov}(D_{0j,p+1}, D_{0jk}) \\ &\quad +2 \sum_{1 \le u < v \le p} \left(\frac{1}{1+\lambda_{0ju}}\right)\left(\frac{1}{1+\lambda_{0jv}}\right) \text{Cov}(D_{0ju}, D_{0jv}) \\ &= \frac{2n_0(n_j-p)}{(n_0+n_j-p)^2(n_0+n_j-p+2)} + \frac{2(n_0+n_j-p-1)}{(n_0+n_j-p)^2(n_0+n_j-p+2)} \sum_{k=1}^{p} \left(\frac{1}{1+\lambda_{0jk}}\right)^2 \\ &\quad - \frac{2(n_j-p)}{(n_0+n_j-p)(n_0+n_j-p+2)} \sum_{k=1}^{p} \left(\frac{1}{1+\lambda_{0jk}}\right) \\ &\quad - \frac{2}{(n_0+n_j-p)(n_0+n_j-p+2)} \sum_{1 \le u < v \le p} \left(\frac{1}{1+\lambda_{0ju}}\right)\left(\frac{1}{1+\lambda_{0jv}}\right). \ (C.6) \end{aligned}$$

Finally, $E(W_{0,j})$ can be calculated by $(C.5)$, $(C.6)$.

For convenient, here calculate $\text{Var}(W_{0,j})$ with the fixed X_0 and X_j case in section 3.

$$X_j \equiv X_{n\times p}, \ X_0 = \begin{pmatrix} X \\ \vdots \\ X \end{pmatrix}_{nm\times p}.$$

Then

$$\begin{aligned}(X_0^T X_0)^{1/2}(X_j^T X_j)^{-1}(X_0^T X_0)^{1/2} &= (mX^T X)^{1/2}(X^T X)^{-1}(mX^T X)^{1/2} \\ &= \frac{1}{m} I_p, \\ \lambda_{0jk} &= \frac{1}{m}, k = 1, 2, \cdots, p.\end{aligned}$$

Hence

$$\begin{aligned}\mathrm{Var}(W_{0,j}) &= \mathrm{Var}\left((m+1)nD_{0j} + \frac{(m+1)^3 n^3}{2mn^2}\left(D_{0j,p+1} + \frac{m}{m+1}D_{0j} - \frac{1}{m+1}\right)^2\right) \\ &= E\left(\left[(m+1)nD_{0j} + \frac{(m+1)^3 n^3}{2mn^2}\left(D_{0j,p+1} + \frac{m}{m+1}D_{0j} - \frac{1}{m+1}\right)^2\right]^2\right) \\ &\quad - \left[E\left((m+1)nD_{0j} + \frac{(m+1)^3 n^3}{2mn^2}\left(D_{0j,p+1} + \frac{m}{m+1}D_{0j} - \frac{1}{m+1}\right)^2\right)\right]^2,\end{aligned}$$

where

$$(D_{0j0},\ \sum_{k=1}^{p} D_{0jk} \equiv D_{0j},\ D_{0j,p+1})^T \sim \mathrm{Dirichlet}\left(\frac{n_0 - p}{2},\ \frac{p}{2},\ \frac{n_j - p}{2}\right),$$

and the first part of $\mathrm{Var}(W_{0,j})$ can be calculated using the previous method.
By the book ≪ *Continuous multivariate distributions* ≫ *p*.488 , we have

$$E\left\{D_{0j}^m D_{0j,p+1}^n\right\} = \frac{(p/2)^{[m]}\left[(n_j - p)/2\right]^{[n]}}{(n_j/2)^{[m+n]}}, \text{ where } k^{[m]} \equiv k(k+1)\cdots(k+m-1).$$

And here can use this method to calculate the second part of $\mathrm{Var}(W_{0,j})$. Finally, we can calculate the $\mathrm{Var}(W_{0,j})$.

D.1

$E_{\boldsymbol{\theta}}(W^*_{0,j})$ and $\mathrm{Var}_{\boldsymbol{\theta}}(W^*_{0,j})$

$$E_{\boldsymbol{\theta}}(W^*_{0,j}) = E_{\boldsymbol{\theta}}(W_{0,j}) - E_{\boldsymbol{\theta}}(W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*_j))$$

where $E_{\boldsymbol{\theta}}(W_{0,j})$ is given in appendix C.1, and by equation (2.19) can know

$$\begin{aligned} & E_{\boldsymbol{\theta}}(W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*_j)) \\ = & \frac{n_0 n_j (n_0 + n_j)}{2} E_{\boldsymbol{\theta}} \left(\left(\frac{1}{n_0} \frac{H_0}{H_0 + H_j} - \frac{1}{n_j} \frac{H_j}{H_0 + H_j} \right)^2 \cdot 1_{\frac{1}{n_j} \frac{H_j}{H_0+H_j} < \frac{1}{n_0} \frac{H_0}{H_0+H_j}} \right) \end{aligned}$$

where $H_0 \sim \chi^2_{n_0 - p}$, $H_j \sim \chi^2_{n_j - p}$ and they are independent. And let

$$B_j \equiv \frac{H_j}{H_0 + H_j} \sim \mathrm{Beta}(\frac{n_j - p}{2}, \frac{n_0 - p}{2})$$

then

$$\begin{aligned} & E_{\boldsymbol{\theta}} \left(\left(\frac{1}{n_0} \frac{H_0}{H_0 + H_j} - \frac{1}{n_j} \frac{H_j}{H_0 + H_j} \right)^2 \cdot 1_{\frac{1}{n_j} \frac{H_j}{H_0+H_j} < \frac{1}{n_0} \frac{H_0}{H_0+H_j}} \right) \\ = & E_{\boldsymbol{\theta}} \left(\left(\frac{1 - B_j}{n_0} - \frac{B_j}{n_j} \right)^2 \cdot 1_{\frac{B_j}{n_j} < \frac{1-B_j}{n_0}} \right) \\ = & E_{\boldsymbol{\theta}} \left(\left[\left(\frac{1}{n_0} + \frac{1}{n_j} \right) B_j - \frac{1}{n_0} \right]^2 \cdot 1_{\left(\frac{1}{n_0} + \frac{1}{n_j} \right) B_j < \frac{1}{n_0}} \right) \\ = & \left(\frac{n_0 + n_j}{n_0 n_j} \right)^2 E_{\boldsymbol{\theta}} \left(\left(B_j - \frac{1/n_0}{1/n_0 + 1/n_j} \right)^2 \cdot 1_{B_j < \frac{1/n_0}{1/n_0 + 1/n_j}} \right), \end{aligned}$$

and let

$$\frac{1/n_0}{1/n_0 + 1/n_j} = \frac{n_j}{n_0 + n_j} \equiv a_j (< 1).$$

Then

$$E_{\boldsymbol{\theta}}(W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*)) = \frac{(n_0+n_j)^3}{2n_0n_j} E_{\boldsymbol{\theta}}\left((B_j - a_j)^2 \cdot 1_{B_j<a_j}\right)$$

and

$$\begin{aligned}
& E_{\boldsymbol{\theta}}\left(B_j^k \cdot 1_{B_j<a_j}\right) \\
= & \int_0^{aj} \frac{x^{k+(n_j-p)/2-1}(1-x)^{(n_0-p)/2-1}}{\Gamma\left((n_j-p)/2\right)\Gamma\left((n_0-p)/2\right)/\Gamma\left((n_0+n_j-2p)/2\right)} dx \\
= & \frac{\Gamma\left((n_0+n_j-2p)/2\right)\Gamma\left(k+(n_j-p)/2\right)}{\Gamma\left((n_j-p)/2\right)\Gamma\left(k+(n_0+n_j-2p)/2\right)} \\
& \cdot \int_0^{aj} \frac{x^{k+(n_j-p)/2-1}(1-x)^{(n_0-p)/2-1}}{\Gamma\left(k+(n_j-p)/2\right)\Gamma\left((n_0-p)/2\right)/\Gamma\left(k+(n_0+n_j-2p)/2\right)} dx \\
= & \frac{\Gamma\left((n_0+n_j-2p)/2\right)\Gamma\left(k+(n_j-p)/2\right)}{\Gamma\left((n_j-p)/2\right)\Gamma\left(k+(n_0+n_j-2p)/2\right)} F(a_j),\ \ k=1,2,3,\cdots,
\end{aligned}$$

where $F(\cdot)$ is the cdf of Beta$(k+(n_j-p)/2, (n_0-p)/2)$. Finally, $E_{\boldsymbol{\theta}}(W_{0,j}^*)$ can be calculated.

$$\mathrm{Var}_{\boldsymbol{\theta}}(W_{0,j}^*) = \mathrm{Var}_{\boldsymbol{\theta}}(W_{0,j}) + \mathrm{Var}_{\boldsymbol{\theta}}(W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*)) - \mathrm{Cov}(W_{0,j}, W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*)).$$

$\mathrm{Var}_{\boldsymbol{\theta}}(W_{0,j})$ is given in appendix C.1, and

$$\mathrm{Var}_{\boldsymbol{\theta}}(W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*)) = E_{\boldsymbol{\theta}}(W_{0,j}^2(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*)) - \left[E_{\boldsymbol{\theta}}(W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*))\right]^2$$

where $E_{\boldsymbol{\theta}}(W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*))$ is calculated in the previous, and

$$E_{\boldsymbol{\theta}}(W_{0,j}^2(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*)) = \frac{(n_0+n_j)^6}{4n_0^2n_j^2} E_{\boldsymbol{\theta}}\left[(B_j - a_j)^4 \cdot 1_{B_j<a_j}\right],$$

it can be calculated using the previous method. $\mathrm{Cov}(W_{0,j}, W_{0,j}(\boldsymbol{\theta}^*, \boldsymbol{\theta}_j^*))$ is too complicated to calculate, so using the simple covariance to estimate it. Finally, $\mathrm{Var}_{\boldsymbol{\theta}}(W_{0,j}^*)$ can be calculated.