AUSTRIAN JOURNAL OF STATISTICS
Volume 27 (1998), Numbers 1&2, 17–37

One-sided Multivariate Testing and Environmental Monitoring

Alessandro Fassò
University of Bergamo, Italy

Abstract: In environmental monitoring the quality assessment is essentially one-sided. For example quality of air decreases as one or more air pollutants increase. Statistical multivariate monitoring is then concerned with the problem of detecting some non-random one-sided shifts of one or more among various pollutant levels.

In this paper, after discussing some change models useful in environmental monitoring, a generalized likelihood ratio test statistic is proposed both for retrospective testing and for on-line change detection. The test is then compared to the approach based on union-intersection of one-sided univariate tests.

Keywords: Order Restricted Inference, One-sided Multivariate *EWMA*, Multivariate Control Charts.

1 Introduction

1.1 Statistical Synthesis of Environmental Monitoring

In analyzing environmental-data collected over time, a typical problem is testing the constancy of some parameter of interest, say θ, which is related to the variability or distribution of the data y. For example θ may be the mean concentration vector of some pollutants y. Two classes of problems may arise, i.e. a) retrospective analysis and b) on-line change detection.

In the first case a), at time t_n, the researcher, on the basis of observed data $y_{t_1}, ..., y_{t_n}$, has to decide for or against the null hypothesis $H_0 : \theta_s = \theta^0$ for $s = 1, .., t_n$. Moreover, he/she can delay the conclusion and continue data collection which can be expensive (see e.g. Shipper et al., 1997).

In the second class of problems b) the data are (automatically) collected at (equally spaced) times $t = 1, 2, ...$. At each time t, using data $y_1, ..., y_t$, a decision has to be taken about the hypothesis $H_0 : \theta_s = \theta^0$ for $s = 1, ..., t$ and if it seems untrue, an *alarm* is to be given.

Under both a) or b) the change in θ, if any, happens at *change time* $t^* \geq 1$ and may be modelled in various ways. Basically we have deterministic changes and stochastic changes. Typical examples of deterministic changes are given by the following three models.

1. Step changes: $\theta(t) = \begin{cases} \theta^0 & t < t^* \\ \theta \neq \theta^0 & t \geq t^* \end{cases}$

2. Linear trends: $\theta(t) = \begin{cases} \theta^0 & t < t^* \\ (t - t^*)\,\theta & t \geq t^* \end{cases}$

3. Smooth transitions: $\theta(t) = \begin{cases} \theta^0 & t < t^* \\ \sigma(t - t^*)\,\theta & t \geq t^* \end{cases}$, $\sigma()$ is $S-$shaped.

Stochastic changes are defined by stochastic equations involving y_t itself, see e.g. Fassò (1997). For example the flow of a basin is highly nonlinear (see e.g. Young and Beven, 1994) and an alarm system based on the above *deterministic* models could be inefficient.

In models above, θ^0 is a point and H_0 is a *simple* hypothesis. Often nuisance parameters arise and/or θ^0 is a set. For example we may want to test for a high concentration of some pollutants so we have the no-pollution hypothesis $H_0 : \theta(t) \leq \theta^0$ for each coordinate, against one sided alternatives $H_+ : \theta(t) > \theta^0$, for some coordinates and $t \geq t^*$. If θ is a scalar the solution of most related problems is simple, but if θ is a vector some care is needed and this will be the concern of the rest of this paper.

The main difference between problems $a)$ and $b)$ above relates to the change time t^* and explains the different statistical procedures used for. In $a)$ the change time t^* may be known or unknown and usually is independent on t, for example often $t^* = 1$. Vice versa in $b)$, the procedure is repeated at each time t and only *recent* change times are of interest, so if t is large $1 << t^* \leq t$.

Hence different performance indices are of interest. For case $a)$ decision error probabilities can be used giving significance level and power. Vice versa for case $b)$ the so-called ARL function can be used giving mean time between false alarms $(MTFA)$ and mean delay to change detection (MTD).

In principle, the two problems could be treated in the same way. Change detection using for example the generalized likelihood ratio (GLR) approach (see e.g. Basseville and Nikiforov, 1993, Section 7.2.1.9) gives tests for H_0 and an estimate for t^*. Despite this in on-line detection problems, the GLR often gives non-recursive algorithms requiring increasing amounts of computing time and memory space. Moreover, in Fassò (1997) there is some evidence showing that this approach is no better than recursive ones in many cases.

Various approaches are available for real time detection which give change-detectors and have graphic representations called control charts from statistical quality control. Among them we have the classical without memory approach of Shewhart, which is more efficient for large changes. Then we have the finite memory approach (e.g. Fassò, 1992), the *CUSUM* approach and the exponentially weighted moving average approach *(EWMA)* which are sensitive to small changes. Often environmental data are autocorrelated and, in this case, a preliminary whitening filter can be used (see e.g. Johnson and Bagshaw, 1974, and Fassò, 1997).

1.2 Aims and Paper Organization

In this paper, approach $a)$ for $t^* = 1$ and approach $b)$ will be considered in some detail for the mean vector θ of multivariate gaussian independent observations. In particular for approach a) a standard test based on the sufficient statistic will be considered and for approach b) both the Shewhart approach and *EWMA* approach will be shortly considered.

A change will be defined by the rise of one or more coordinates of θ. This is relevant in most environmental monitoring and safety problems where the decrease of e.g. air pollution or water flow is not a matter of alarm systems. Of course, sometimes a decrease in some coordinates may be of interest; e.g. O_2 in air quality monitoring. In this case, simply changing the sign of the corresponding deviates $y_{tj} - \theta_j^0$ is the way to use the results of next sections.

The standard approach to multivariate level testing and on-line monitoring is based on the Hotelling's χ^2 or T^2 statistic. This statistic is direction invariant and if used for our problem would give a lot of false alarms due to decrease of the parameters. Moreover there is a loss in efficiency due to the miss-specification of the hypotheses.

In Fassò (1996), I considered the case of a simple non-change hypothesis $\theta(t) = \theta^0$ obtaining non-standard statistics and tests. I proposed a non-symmetric multivariate *EWMA* detector. Using Monte Carlo experiments, I compared the new one-sided multivariate to the linear *CUSUM* detector and to the standard quadratic *MEWMA* of Lowry et al. (1992) giving figures for the improvement in mean delay for the new method.

In the following, a non-symmetric statistic will be developed for testing the complex null hypothesis of non-increase defined in Section 1.3. This problem is a particular case of the so-called *order restricted inference* which has received a lot of attention in the past. In particular, Kudo (1963) and Nuesch (1966) considered the closely related problem of Fassò (1996) mentioned above. Moreover, for recent generalizations and a review, see, e.g. Robertson et al. (1988) or Shapiro (1988). Here the application to environmental monitoring and control chart seems to be new. Further the new expressions of the GLR statistics and its distribution seems to improve computability in certain respects and the new results of Section 2.2 are of some practical interest.

As a matter of fact this statistic is here proposed both as a standard test statistic and as a basis of a Shewhart or *MEWMA* detector.

The paper is organized as follows. In Section 1.3 the appropriate change model is defined and some symbols are defined. In Section 2 considering the simple bivariate case, the restricted maximum likelihood (MLE) estimate is used to get the GLR statistic and its distribution. In Section 2.2 the expectation is given in closed form and used to propose a test statistic almost independent on the correlation ρ.

In Section 2.3 on-line change detection or monitoring is considered using Shewhart or *EWMA* detectors applied to the introduced GLR statistic.

In Section 3 the new test statistic is compared with the simple procedure based on separately testing each coordinate. In Section 4 the results are extended to the general multivariate case giving computable formulas to implement the GLR statistic and its distribution. Finally, in Section 5, some concluding remarks are given.

1.3 The One Sided Change Model

Suppose the monitored data are independent $k-$variate normal, i.e. y_t, observed at times $t = 1, 2, \dots$, are $IIN_k(\theta(t), \Sigma)$, with mean vector $\theta(t) = (\theta_1(t), \dots, \theta_k(t))' \in \Re^k$ and covariance matrix $\Sigma = (\rho_{ij}) > 0$ which is supposed constant and known with $\rho_{ii} = 1$.

Suppose θ^0 is known and put $\theta^0 = 0$, then the *no-increase* parameter sub-space is given by

$$\Theta_0 = \left\{\theta \in \Re^k : \theta_j \leq 0, j = 1, ..., k\right\}.$$

Correspondingly the *increase* parameter sub-space is given by

$$\Theta_1 = \Re^k - \Theta_0 = \left\{\theta \in \Re^k : \theta_j > 0, \text{ for some } j = 1, ..., k\right\}.$$

With these symbols the *no-increase hypothesis* is given by

$$H_0 : \theta(t) \in \Theta_0, \ t = 1, 2, ...$$

and the *increase hypothesis* is

$$H_1 = \begin{cases} \theta(t) \in \Theta_0 & t < t^* \\ \theta(t) \in \Theta_1 & \text{else} \end{cases}$$

where $t^* = 1$ for standard retrospective testing and t^* is unknown for on-line detection.

Now, letting for simplicity $k = 2$, consider the statistic given in Fassò (1996), say U, which is a GLR test statistic for the null hypothesis $\theta = 0$ against one-sided alternatives $\theta_j \geq 0$. It is easy to see that, using this test statistic, for the new problem of complex null hypothesis, it gives a test which is biased against $\theta \in \Theta_0$ away from the diagonal $\theta_1 = \theta_2$. To be more precise, using as a critical value u_α

$$P(U > u_\alpha; \rho, \theta = 0) = \alpha$$

we get a biased test whose power exceeds α for various $\theta's$ belonging to Θ_0.

2 The Bivariate GLR Approach

In order to get a test statistic of size not exceeding α for any $\theta \in \Theta_0$, let $y = (y_1, y_2)$ have the bivariate normal distribution, $N_2(\theta, \Sigma)$, with density $f(y; \theta)$ say, and $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ with $\rho^2 < 1$.

Now consider the GLR statistic

$$Q(y) = -2 \log \left(\frac{f\left(y; \hat{\theta}\right)}{f(y; y)} \right)$$

where $\hat{\theta}$ is the MLE restricted to Θ_0 given in Section 2.1 below and, of course, y is the unrestricted MLE of θ. This statistic can be viewed as the estimated Rao's score test statistic (see e.g. Conniffe, 1990), i.e.

$$Q(y) = l\left(\hat{\theta}\right)' \Sigma l\left(\hat{\theta}\right)$$

where

$$l(\theta) = \frac{\partial}{\partial \theta} \log f(y; \theta)$$

and for the bivariate gaussian distribution, it gives

$$Q(y) = \left(y - \hat{\theta}\right)' \Sigma^{-1} \left(y - \hat{\theta}\right). \quad (1)$$

2.1 The Restricted MLE

The restricted MLE of $\theta \in \Theta_0$ based on the data y, say $\hat{\theta} = \hat{\theta}(y)$, is the projection of y into Θ_0 minimizing $(y-\theta)' \Sigma^{-1} (y-\theta)$ for $\theta \in \Theta_0$. It is easy to see (see e.g. Fassò, 1996) that

$$\hat{\theta}(y) = yI_{(0)} + \hat{\theta}_{(1)} I_{(1)} + \hat{\theta}_{(2)} I_{(2)} \tag{2}$$

where

$$\begin{aligned} I_{(i)} &= \left\{ \begin{array}{ll} 1 & y \in \Theta_i \\ 0 & else \end{array} \right. \quad i = 0, .., 3 \\ \Theta_1 &= \{y_1 > 0,\ y_2 < \rho y_1\} \\ \Theta_2 &= \{y_2 > 0,\ y_1 < \rho y_2\} \\ \Theta_3 &= \Re^2 - \Theta_0 - \Theta_1 - \Theta_2 = \{y_1 > \rho y_2,\ y_2 > \rho y_1\} \end{aligned} \tag{3}$$

and

$$\begin{aligned} \hat{\theta}_{(1)} &= \left(\begin{array}{c} 0 \\ y_2 - \rho y_1 \end{array} \right), \\ \hat{\theta}_{(2)} &= \left(\begin{array}{c} y_1 - \rho y_2 \\ 0 \end{array} \right). \end{aligned}$$

Note that $\hat{\theta}(y) = 0$ for $y \in \Theta_3$.

2.2 The GLR Statistic

In order to get the statistic Q, note that $I_{(i)} I_{(j)} = 0$ with probability one and

$$y - \hat{\theta} = y_1 \left(\begin{array}{c} 1 \\ \rho \end{array} \right) I_{(1)} + y_2 \left(\begin{array}{c} \rho \\ 1 \end{array} \right) I_{(2)} + yI_{(3)}.$$

Then, from (1), we have the estimated score

$$Q(y) = y_1^2 I_{(1)} + y_2^2 I_{(2)} + y' \Sigma^{-1} y I_{(3)} \tag{4}$$

with large values giving empirical evidence against the no-increase hypothesis.

In order to evaluate significance and power of the related test, the cumulative distribution function (cdf) of Q is given in computable form in the following proposition which uses some integrals involving the standard normal density and *cdf* denoted respectively by $\phi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ and $\Phi(t) = \int_{-\infty}^{t} \phi(x)\, dx$. In particular, the integrals

$$M(\mu, \eta)_a^b = \int_a^b \Phi(\mu + \eta t)\, \phi(t)\, dt$$

and

$$N(\mu, \eta, h)_a^b = \int_a^b \Phi\left(\mu + \sqrt{h^2 - (t+\eta)^2}\right) \phi(t)\, dt$$

are used in the sequel and, from the practical point of view, can be evaluated by numeric quadrature. Moreover, let $\Phi_{\mu,\Sigma}(t)$ with $t = (t_1, ..., t_k)'$ be the $k-$variate normal cdf, $N_k(\mu, \Sigma)$, and $\Phi_\Sigma = \Phi_{0,\Sigma}$.

Then the following proposition is proven in the Appendix using $F_{\chi^2_d}$ for the cdf of the χ^2 distribution with d degrees of freedom.

Proposition 1 *For* $t \geq 0$ *we have*

$$\begin{aligned} F_Q\left(t^2;\rho,\theta\right) &= P\left(Q \leq t^2\right) \\ &= M\left(\frac{-\theta_2}{\sqrt{1-\rho^2}}, \frac{-\rho}{\sqrt{1-\rho^2}}\right)_{-\infty}^{-\theta_1} + \Phi\left(\frac{\rho\theta_2-\theta_1}{\sqrt{1-\rho^2}}\right)\left(\Phi\left(t-\theta_2\right)-\Phi\left(-\theta_2\right)\right) \\ &\quad + M\left(\frac{\theta_1-\rho\theta_2}{\rho\sqrt{1-\rho^2}}, \frac{\sqrt{1-\rho^2}}{\rho}\right)_{-\theta_1}^{\rho t-\theta_1} + N\left(\frac{\rho\theta_1-\theta_2}{\sqrt{1-\rho^2}}, \theta_1, t\right)_{\rho t-\theta_1}^{t-\theta_1}. \end{aligned} \quad (5)$$

In particular, when $\theta = 0$, *we have*

$$\begin{aligned} F_Q\left(t^2;\rho\right) &= F_Q\left(t^2;\rho,0\right) \\ &= \Phi(t) + F_{\chi^2_1}(t)\frac{\arctan\sqrt{\frac{1-\rho}{1+\rho}}}{\pi} + \frac{1}{2\pi}\arcsin(\rho) - \frac{1}{4}. \end{aligned} \quad (6)$$

Note that Q *has a probability mass on the origin given by*

$$P(Q=0) = M\left(\frac{-\theta_2}{\sqrt{1-\rho^2}}, \frac{-\rho}{\sqrt{1-\rho^2}}\right)_{-\infty}^{-\theta_1}$$

and, when $\theta = 0$, *this is the so-called orthant probability*

$$P(Q=0) = \frac{1}{4} + \frac{1}{2\pi}\arcsin(\rho).$$

Some quantiles q_α of this distribution for $\theta = 0$ are reported in Table 1.

In some cases, it may be interesting to reduce the influence of ρ on F_Q. If the expectation $E(Q;\rho,\theta)$ is known, one can then use the centered statistic $Q - E(Q)$. This is relevant for example when ρ is estimated or only approximately known or not exactly constant from observation to observation. To get computable expressions for $E(Q)$ let us introduce the following integrals.

Definition 1 *Let*

$$K_{m,a}^b = \int_a^b t^m \phi(t)\,dt, \quad (7)$$

$$L_m(\mu,\eta)_a^b = \int_a^b t^m \phi(\mu+\eta t)\,\phi(t)\,dt, \quad (8)$$

and

$$M_m(\mu,\eta)_a^b = \int_a^b t^m \Phi(\mu+\eta t)\,\phi(t)\,dt\,. \quad (9)$$

Note that, with this notation,

$$M\left(\mu,\eta\right)_a^b = M_0\left(\mu,\eta\right)_a^b \ .$$

Moreover, the integrals K, L and M could be computed by direct quadrature. Nevertheless, in order to reduce numerical integration the following proposition gives useful recursions which can be easily proven using standard integration by parts.

Proposition 2 *Let $m > 1$, then the following recursions hold.*

1. *Let $K_m\left(t\right)$ be the indefinite integral corresponding to (7), then $K_0\left(t\right) = \Phi\left(t\right)$, $K_1\left(t\right) = -\phi\left(t\right)$ and*

$$K_m\left(t\right) = \left(m-1\right)K_{m-2}\left(t\right) - t^{m-1}\phi\left(t\right).$$

2. *Let $L_m\left(t\right)$ be the indefinite integral corresponding to (8), then*

$$L_0\left(t\right) = \frac{1}{\sqrt{\eta^2+1}}\phi\left(\frac{\mu}{\sqrt{\eta^2+1}}\right)\Phi\left(\sqrt{\eta^2+1}\left(t+\frac{\mu\eta}{\eta^2+1}\right)\right)$$

and

$$L_1\left(t\right) = -\frac{1}{\eta^2+1}\phi\left(\frac{\mu}{\sqrt{\eta^2+1}}\right)$$
$$\times\left(\phi\left(\sqrt{\eta^2+1}\left(t+\frac{\mu\eta}{\eta^2+1}\right)\right)+\frac{\mu\eta}{\sqrt{\eta^2+1}}\Phi\left(\sqrt{\eta^2+1}\left(t+\frac{\mu\eta}{\eta^2+1}\right)\right)\right).$$

Moreover, if $\eta \neq -1$,

$$L_m\left(t\right) = \frac{\left(m-1\right)L_{m-2}\left(t\right) - t^{m-1}\phi\left(\mu+\eta t\right)\phi\left(t\right)}{\eta+1}$$

and, if $\eta = -1$,

$$L_m\left(t\right) = -\frac{t^{m+1}}{m+1}\phi\left(t-\mu\right)\phi\left(t\right).$$

3. *Let $M_m\left(t\right)$ be the indefinite integral corresponding to (9), then*

$$M_1\left(t\right) = \eta L_0\left(t\right) - \phi\left(t\right)\Phi\left(\mu+\eta t\right)$$

and

$$M_m\left(t\right) = \left(m-1\right)M_{m-2}\left(t\right) + \eta L_{m-1}\left(t\right) - t^{m-1}\Phi\left(\mu+\eta t\right)\phi\left(t\right).$$

As a consequence of this proposition, only M_0 has to be computed using numerical integration. For example, the following relations are useful:

1. $K^{\infty}_{2,-\theta} = \Phi(\theta) - \theta\phi(\theta)$,
2. $K^{\infty}_{3,-\theta} = (2+\theta^2)\,\phi(\theta)$,
3. $K^{\infty}_{4,-\theta} = 3\Phi(\theta) - \theta(3+\theta^2)\,\phi(\theta)$,
4. $K^{b}_{m,-\infty} = (-1)^m\, K^{\infty}_{m,-\theta}$,
5. $L^{b}_{m,a} = L^{-b}_{m,-a}$.

Using these, in the Appendix is proven the following proposition which gives computable formulas for the expectation of the statistic Q.

Proposition 3

$$E(Q|\theta) = 2 + \theta'\Sigma^{-1}\theta - \frac{1}{1-\rho^2}\sum_{s=0}^{2}\left(\sum_{i=1,2} E\left(y_i^2 I_{(s)}\right)\rho^{2I(i=s)} - 2\rho E\left(y_1 y_2 I_{(s)}\right)\right)$$

where, for $i = 1, 2$, *we have*

1. $E\left(y_i^2 I_{(i)}\right) = \Phi\left(\frac{\rho\theta_i - \theta_{3-i}}{\sqrt{1-\rho^2}}\right)\left(\left(1+\theta_i^2\right)\Phi(\theta_i) + \theta_i\phi(\theta_i)\right)$,
2. $E\left(y_i^2 I_{(3-i)}\right) = \left(1-\rho^2\right) K^{+\infty}_{2,\beta_i}\Phi(\theta_{3-i})$

   $+\Phi\left(\frac{\rho\theta_{3-i} - \theta_i}{\sqrt{1-\rho^2}}\right)\left(\theta_i^2\Phi(\theta_{3-i}) + \rho^2 K^{+\infty}_{2,-\theta_{3-i}} + 2\rho\theta_i\phi(\theta_{3-i})\right)$

   $-2\sqrt{1-\rho^2}\,\phi\left(\frac{\theta_i - \rho\theta_{3-i}}{\sqrt{1-\rho^2}}\right)\left(\theta_i\Phi(\theta_{3-i}) + \rho\phi(\theta_{3-i})\right)$,
3. $E\left(y_i^2 I_{(0)}\right) = M^{\theta_i}_{2,-\infty}(\mu_{3-i},\eta) + 2\theta_i M^{\theta_i}_{1,-\infty}(\mu_{3-i},\eta) + \theta_i^2 M^{\theta_i}_{0,-\infty}(\mu_{3-i},\eta)$,
4. $E\left(y_1 y_2 I_{(0)}\right) = (\theta_1 + \rho\theta_2)\, M^{\theta_2}_{1,-\infty}(\mu_1,\eta) + \rho M^{\theta_2}_{2,-\infty}(\mu_1,\eta)$

   $+\theta_1\theta_2 M^{\theta_2}_{0,-\infty}(\mu_1,\eta) - \sqrt{1-\rho^2}\left(L^{\theta_2}_{1,-\infty}(\mu_1,\eta) + \theta_2 L^{\theta_2}_{0,-\infty}(\mu_1,\eta)\right)$,
5. $E\left(y_i y_{3-i} I_{(i)}\right) = -\sqrt{1-\rho^2}\,\phi\left(\frac{\theta_{3-i} - \rho\theta_i}{\sqrt{1-\rho^2}}\right)\left(\theta_i\Phi(\theta_i) + \phi(\theta_i)\right)$

   $+\Phi\left(\frac{\rho\theta_i - \theta_{3-i}}{\sqrt{1-\rho^2}}\right)\left(\theta_{3-i}\phi(\theta_i) + \theta_i\theta_{3-i}\Phi(\theta_i) + \rho\Phi(\theta_i)\right)$

   where $\beta_i = \frac{\theta_i - \rho\theta_{3-i}}{\sqrt{1-\rho^2}}$, $\mu_i = \frac{-\theta_i}{\sqrt{1-\rho^2}}$ *and* $\eta = \frac{-\rho}{\sqrt{1-\rho^2}}$.

Table 1 gives some expected values of F_Q which may be useful in practice.

Now, following e.g. Conniffe (1990), we can get an approximately invariant procedure for the complex H_0 using

$$Q - E(Q;\rho,\theta=0)$$

or

$$Q - E\left(Q;\rho,\theta=\hat{\theta}\right).$$

	$1-\alpha$				
ρ	0.90	0.95	0.99	0.995	$E(Q)$
-0.9	3.594	4.915	8.035	9.392	1.356
-0.5	3.275	4.577	7.671	9.021	1.167
0	2.952	4.231	7.289	8.628	1
0.5	2.580	3.820	6.823	8.144	0.8333
0.9	2.080	3.245	6.129	7.413	0.6436

Table 1: Quantiles h_α and expected values of F_Q for $\theta = 0$

2.3 Tests and Detectors Based on Q

If we are interested in testing H_0 against H_1 for $t^* = 1$ (see Section 1.3) using $y_1, ..., y_n$, from the previous section, we will use large values of the test statistic

$$Q\left(y_n^*\right)$$

where $y_n^* = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} y_t$ is gaussian $N(\theta, \sum)$. For example rejecting H_0 when $Q\left(y_n^*\right) > q_\alpha$, where q_α is the corresponding percentile of Q for $\theta = 0$, gives an $\alpha - sized$ critical region. Note that the convex shape of both Θ_0 and the contours of the power function $\pi\left(\theta; \rho\right) = 1 - F_Q\left(q_\alpha; \rho, \theta\right)$ (see Figures $1a$ and b to $3a$ and b) gives rise to a small bias along the lines $\theta_1 = 0$ and $\theta_2 = 0$. Fortunately, the *bias region, i.e.* the subset of Θ_1 where $\pi\left(\theta; \rho\right) < \alpha$, drawn in Figures 1d to 3d (using $\alpha = 0.05$), is not large and, by continuity, the bias too, is small.

Turning to problem b) of the introduction, if we are interested in on-line monitoring, we can use a Shewhart-like control chart which is suitable for large shifts in θ. This gives a warning or an alarm at time t if

$$Q\left(y_t\right) > q_\alpha.$$

In this case, the mean time between false alarms is *MTFA*$= \frac{1}{\alpha}$ and the mean delay is

$$MTD\left(\theta\right) = \frac{1}{\pi_Q\left(\rho, \theta\right)}.$$

Hence the detector's performances can be assessed using proposition 1.

Vice versa, a detector sensitive to small-to-moderate changes can be obtained using the *EWMA* transform, i.e.

$$z_t = \lambda y_t + \left(1 - \lambda\right) z_{t-1}$$

with $z_0 = 0$ and $0 < \lambda < 1$ ($\lambda = 0.1$ is often used in quality control). This gives the one-sided *EWMA* procedure which declares alarm at time t if

$$Q\left(z_t\right) > h.$$

The constant h has to be chosen to protect both from too often false alarms and too late true alarms. In practice, the threshold h can be estimated using Monte Carlo simulations, for example as in Fassò (1996), or using stochastic approximation as in Masarotto and Capizzi (1998). The classical control chart approach then charts $Q(y_t)$ or $Q(z_t)$ against t allowing graphical assessment of possible shifts. For a recent discussion on multivariate control chart see e.g. Mason et al. (1996).

In Section 4, the multivariate extension of the statistic Q is considered. Whenever Q is given in closed form, formulas for its cdf or moments become computationally and formally more and more complex (as a matter of fact, even the general multivariate normal integrals are not easy to compute, see e.g. Johnson and Kotz, 1972). The ARL function of multivariate *EWMA* is not easily computed. For example the approach of Runger and Prabhu (1996) using direction invariance of Hotelling's χ^2 statistic allows dimension reduction to bivariate Markov chains. Here, of course, we cannot use direction invariant statistics as our problem is not direction invariant. Because of this we suggest Monte Carlo-based threshold estimates which of course depend on Σ.

3 The Double Univariate One-sided Approach

The GLR approach of the previous section is relatively complicated for the bivariate problem and, as will be clear in the next section, is more complicated for the general multivariate case. Then it is interesting to know if there are some short cuts retaining the one-sided nature of the problem and giving simpler statistics. In Fassò (1996) it is shown that a detector based on large values of y_1+y_2 has good performances for $\theta_1 \cong \theta_2$ but very low performances for $\theta_1 + \theta_2$ small, even far away from the origin. A simple alternative is to test each coordinate y_i separately and, using e.g. a union-intersection approach, have a warning if at least one of the one-sided univariate tests exceeds an appropriate critical value. This is the same as considering the statistic

$$R = max\left(max\left(0, y_1\right), max\left(0, y_2\right)\right)$$

with distribution function given by

$$F_R\left(t; \rho, \theta\right) = P\left(y_1 \leq t, y_2 \leq t; \rho, \theta\right) = M\left(\frac{t-\theta_2}{\sqrt{1-\rho^2}}, \frac{-\rho}{\sqrt{1-\rho^2}}\right)_{-\infty}^{t-\theta_1}.$$

An α-size test will reject the null hypothesis of Section 1.3 if

$$R > r_\alpha$$

where

$$F_R\left(r_\alpha; \rho, \theta = 0\right) = 1 - \alpha.$$

In order to compare the performances of these statistics we consider some power related quantities. In Figures 1 to 3, for $J = Q$ and R, curves $\pi_J\left(\theta\right) = \alpha$, defining the bias regions, contour plots of the power surfaces $\pi_J\left(\theta\right)$ and contour plots of the relative efficiency surfaces $eff_{Q|R}\left(\theta\right) = \frac{\pi_Q(\theta)}{\pi_R(\theta)}$ are drawn. Some considerations follow.

1. The bias region of Q is very closed to the corresponding region of R, for any ρ (see Figure 1d to 3d where solid and dashed lines overlap). It is generally small and never overlaps the positive quadrant nor exceeds the line $\theta_2 = -\theta_1$.

2. For large positive correlations ($\rho = 0.75$, Figures 1a to 1c) the tests on Q and R are almost equivalent having approximately the same power.

3. When $\rho = 0$, the statistic Q is better than R for positive changes (see Figure 2a to 2c). In this case, the gain in efficiency may achieve 15% being $Eff_{Q|R} = 1.152$ for $\theta = (1.2\;1.2)$. For changes above a line near to $\theta_2 = -\theta_1$, we have $eff_{Q|R} > 0.9$.

4. For large negative correlations ($\theta = -0.75$, Figure 3) the statistic Q is better than R for a large set above the line $\theta_2 = -\theta_1$. In this case the efficiency may be very high, for example $Eff_{Q|R} = 2.53$ for $\theta = (0.8\;0.8)'$.

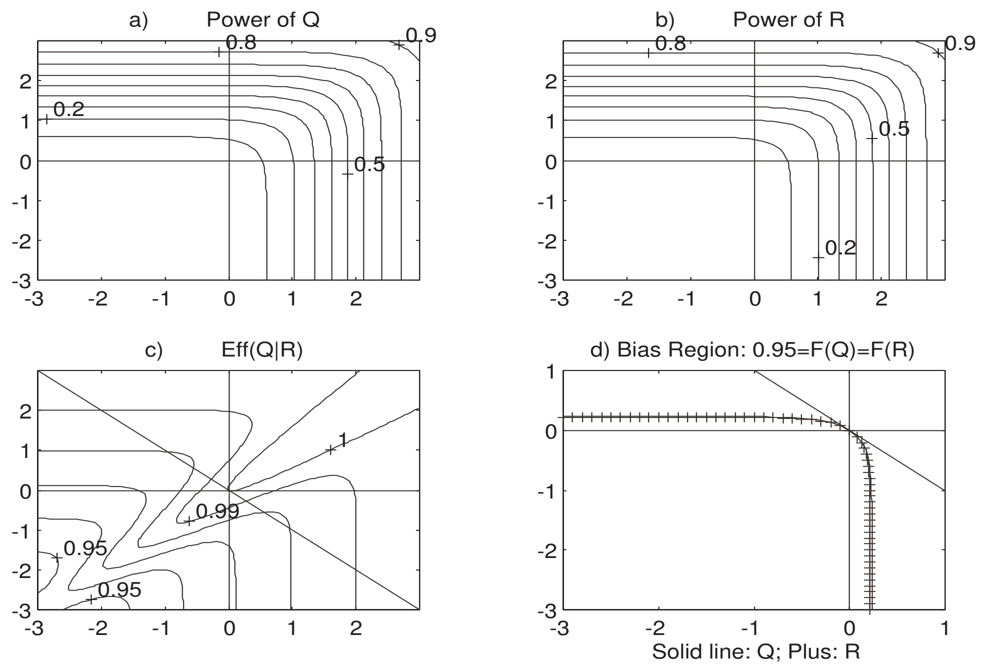

Figure 1: Contours for $\rho = 0.75$; plane (θ_1, θ_2).

4 The Multivariate Case

In this section the approach of Section 2 is generalized to cope with multidimensional data $y = (y_1, ..., y_k)'$. In order to give statistics and $cdf's$ in closed form some more notations are needed.

Firstly in the following two definitions, some permutation operators are introduced.

Definition 2 *Let*

$$\pi = (\pi_1, ..., \pi_k)$$

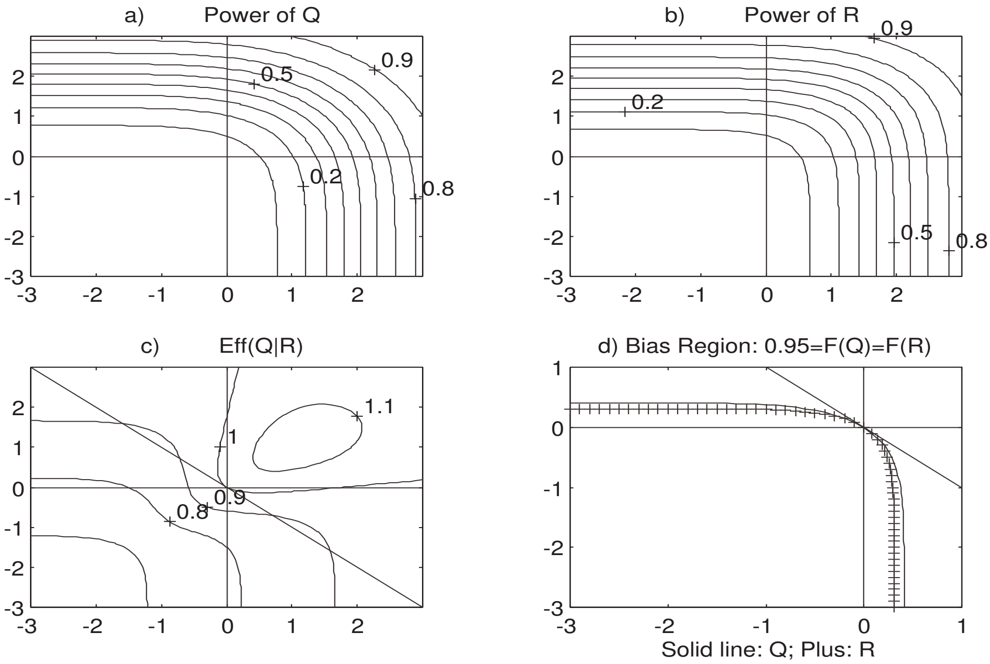

Figure 2: Contours for $\rho = 0$; plane (θ_1, θ_2).

be a permutation of integers $(1, ..., k)$. *Correspondingly, for a column vector* $y = (y_1, .., y_k)'$ *let*

$$x = \pi(y) = (y_{\pi_1}, ..., y_{\pi_k})',$$

and, for a row vector y', *let*

$$\pi(y') = (y_{\pi_1}, ..., y_{\pi_k}).$$

Moreover let π^{-1} *the reverse permutation be such that*

$$y = \pi^{-1}(\pi(y)).$$

Definition 3 *For a permutation* $\pi = (\pi_1, ..., \pi_k)$ *and a square matrix* $A = (a_{ij})$, *let*

$$\pi(A) = \left(a_{\pi_i \pi_j}\right)$$

and

$$A = \pi^{-1}(\pi(A)).$$

Now, in order to define the projection $\hat{\theta} \in \Theta_0$ note that, since Θ_0 is convex, $\hat{\theta}$ is unique. Moreover, following Fassò (1996), consider the frontier

$$\bar{\Theta}_0 = \{\theta \in \Theta_0 : \theta_j = 0 \text{ for some } j = 1, ..., k\}$$

and its partition into the following $p = 2^k - 1$ hyperplanes, R_s, $s = 1, ..., p$, defined using permutations $\pi_s = (\pi_{s1}, ..., \pi_{sk})$, i.e.

$$R_s = \{\theta \in \Theta_0 : \theta_{\pi_{si}} = 0, i = 1, ..., q_s \text{ and } \theta_{\pi_{si}} < 0, i = q_s + 1, ..., k\}.$$

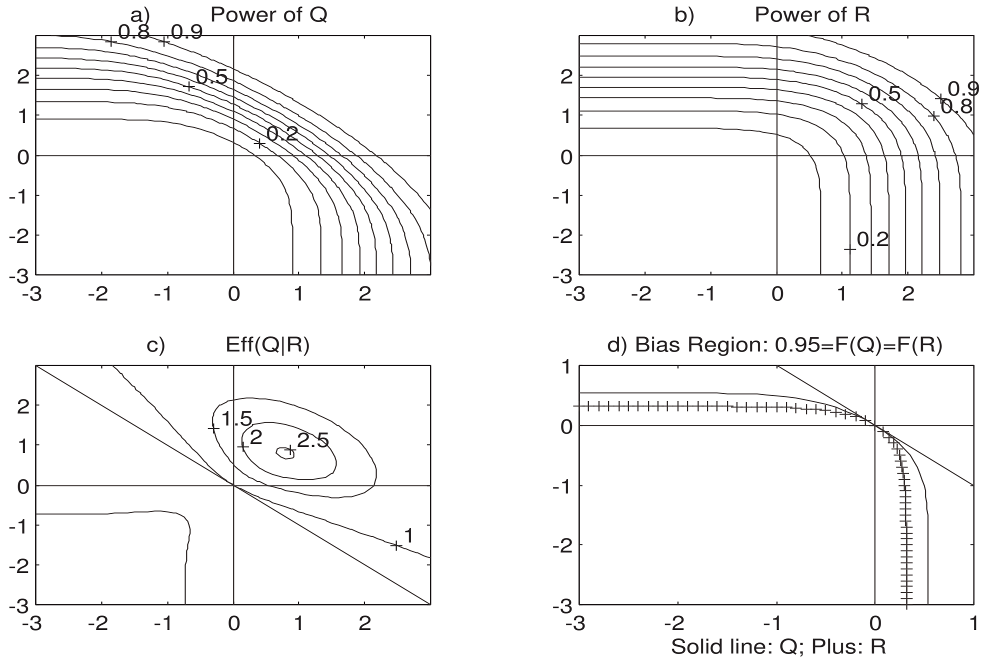

Figure 3: Contours for $\rho = -0.75$; plane (θ_1, θ_2) .

Definition 4 *For each permutation* π_s, $s = 1, ..., p$, *let*

$$\eta_{(s)} = \pi_s(\theta) = (\theta_{\pi_{s1}}, ..., \theta_{\pi_{sk}})'$$

for $\theta \in \Re^k$.

In the sequel, the vector $\eta_{(s)}$ will be partitioned according to first q_s coordinates which are zero for $\theta \in R_s$, i.e.

$$\eta_{(s)} = \left(\theta_{\pi_{s1}}, ..., \theta_{\pi_{sq_s}}, \theta_{\pi_{sq_s+1}}, ..., \theta_{\pi_{sk}}\right)' = \left(\alpha'_{(s)}, \beta'_{(s)}\right)'$$

say, where $\alpha_{(s)} \in \Re^{q_s}$.

Now, let the vector

$$x_{(s)} = \pi_s(y)$$

and its covariance

$$\Delta_{(s)} = \pi(\Sigma) = E\left(x_{(s)} x'_{(s)}\right)$$

be conformably partitioned, i.e.

$$x_{(s)} = \pi_s(y) = \left(a'_{(s)}, b'_{(s)}\right)',$$

where $a_{(s)} = \left(a_{(s),1}, ..., a_{(s),q_s}\right)'$ and

$$\Delta_{(s)} = \begin{pmatrix} \Delta_{(s)aa} & \Delta_{(s)ab} \\ \Delta_{(s)ba} & \Delta_{(s)bb} \end{pmatrix}.$$

4.1 The GLR Statistic

We are going to consider the projection of y into $\bar{\Theta}_0$ in closed form. To do this we use $0_q = (0, ..., 0)'$ for the origin in $\Re^q$, $I()$ for the indicator function and $[f(\cdot)]_i$ for the i^{th} coordinate of the vector valued function $f(\cdot)$.

Now, using the above defined notations, for permutations π_s, $s = 1, ..., p$, consider the partial projections

$$\hat{\eta}_{(s)} = \begin{pmatrix} 0_{q_s} \\ b_{(s)} - \Delta_{(s)ba}\Delta_{(s)aa}^{-1}a_{(s)} \end{pmatrix}$$

and the projecting subset indicators

$$\begin{aligned} I_{(0)}(y) &= I(y \in \Theta_0), \\ I_{(s)}(y) &= I\Big(a_{(s)j} > 0,\ j = 1, ..., q_s, \\ &\qquad \left[b_{(s)} - \Delta_{(s)ba}\Delta_{(s)aa}^{-1}a_{(s)}\right]_j < 0,\ j = 1, ..., k - q_s\Big), \end{aligned} \tag{10}$$

for $s = 1, ..., p-1$, and

$$I_{(p)}(y) = 1 - \sum_{s=0}^{p-1} I_{(s)}(y).$$

Note that $\hat{\eta}_{(p)} = 0_k$. Hence we have the following proposition, which is proven in the Appendix and gives the constrained MLE and the GLR statistic.

Proposition 4 *The restricted MLE minimizing (1) for $\hat{\theta} \in \Theta_0$ is given by*

$$\hat{\theta} = yI_{(0)} + \sum_{s=1}^{p} \pi_s^{-1}\left(\hat{\eta}_{(s)}\right) I_{(s)}. \tag{11}$$

Moreover

$$Q = \sum_{s=1}^{p} Q_{(s)}I_{(s)}, \tag{12}$$

where, for $s = 1, ..., p-1$,

$$Q_{(s)} = a'_{(s)}\Delta_{(s)aa}^{-1}a_{(s)}$$

and

$$Q_{(p)} = y'\Sigma^{-1}y.$$

Now, in order to write the *cdf* of Q for $\theta = 0$, let $S_q(t)$ be the q-dimensional sphere of radius t, i.e.

$$S_q(t) = \left\{w = (w_1, ..., w_q)' : w'w \le t^2\right\},$$

with volume

$$Vol\left(S_q\left(t\right)\right) = \frac{t^q \pi^{q/2}}{\frac{q}{2}\Gamma\left(q/2\right)}.$$

Moreover, let $A_{(0)}\left(t\right)$, be the intersection of the sphere $S_k\left(t\right)$ and the negative cone as follows.

$$A_{(0)}\left(t\right) = \left\{w = \left(w_1, ..., w_k\right)' : \ w'w \leq t^2, \ \Sigma^{1/2} w < 0\right\},$$

let $A_{(s)}\left(t\right)$, for $s = 1, ..., p-1$, be the intersection of the sphere $S_{q_s}\left(t\right)$ and a positive cone as follows.

$$A_{(s)}\left(t\right) = \left\{w = \left(w_1, ..., w_{q_s}\right)' : \ w'w \leq t^2, \ \Delta_{(s)aa}^{1/2} w > 0\right\},$$

and let

$$A_{(p)}\left(t\right) = \left\{w = \left(w_1, ..., w_k\right)' : \ w'w \leq t^2, \ I_{(p)}\left(\Sigma^{1/2} w\right) = 1\right\}.$$

Then $P\left(Q \leq t^2; \Sigma, \theta = 0\right)$ is given in the following proposition which is proven in the Appendix.

Proposition 5 *Let*

$$w_s = \Delta_{(s)aa}^{-1/2} a_s$$

and

$$\Delta\left(w_s\right) = \Delta_{(s)bb} - \Delta_{(s)bw} \Delta_{(s)bb}^{-1} \Delta_{(s)wb}.$$

Then

$$\begin{aligned} F_Q\left(t^2\right) =& \Phi_\Sigma\left(0_k\right) + \sum_{s=1}^{p-1} \Phi_{\Delta(w_s)}\left(0_{q_s}\right) P\left(\chi_{q_s}^2 \leq t^2\right) \frac{Vol\left(A_{(s)}\left(t\right)\right)}{Vol\left(S_{q_s}\left(t\right)\right)} \\ &+ P\left(\chi_k^2 \leq t^2\right) \frac{Vol\left(A_{(p)}\left(t\right)\right)}{Vol\left(S_k\left(t\right)\right)} \end{aligned}$$

where

$$Vol\left(A_{(p)}\left(t\right)\right) = Vol\left(S_k\left(t\right)\right) - \sum_{s=0}^{p-1} Vol\left(A_{(s)}\left(t\right)\right).$$

5 Conclusions

From above examples and considerations, it is evident that none of the statistics considered is always better than the other. Due to the shape of the parameter space under the hypotheses at hand, it is hard to think that such a statistic do exist at all.

Keeping this in mind, we may note that the approach proposed improves both the classical direction-invariance based χ^2 approach which would give undue false alarms and would lack in efficiency and a simple linear combination approach which is optimal only in a direction.

We can conclude that an ad hoc approach is indicated. For the bivariate case this paper gives full results. If changes inside or near the positive quadrant are of interest, it seems that the new GLR is globally superior than the multiple univariate approach. Nevertheless, the additional computations required for the new statistic may be avoided for highly positively correlated observations by using the multiple univariate one-sided approach.

For more general multivariate problems Monte Carlo and/or Bootstrap analysis may help in computing critical values and performance indices.

Acknowledgments

I wish to thank Salvatore Vassallo and Carla Peri for encouraging geometrical considerations.

6 Appendix

In this section, the computations to get the propositions of previous sections are briefly sketched.

6.1 Proof of Proposition 1

First note that

$$P\left(Q=0\right)=P\left(y_i\leq 0, i=1,2\right)=\Phi_{\theta,\Sigma}\left(0\right)$$

and, using the conditional distribution of $y_2|y_1$ for this double integral, we get

$$P\left(Q=0\right)=M\left(\frac{-\theta_2}{\sqrt{1-\rho^2}},\frac{-\rho}{\sqrt{1-\rho^2}}\right)_{-\infty}^{-\theta_1}=p_0\ .$$

Next from (4), we have

$$\begin{aligned}F_Q\left(t^2\right)&=p_0+\sum_{i=1,2}P\left(y_i<t,I_{(i)}=1\right)+P\left(y'\Sigma^{-1}y<t^2,I_{(3)}=1\right)\\&=\sum_{i=0}^{3}p_i\ .\end{aligned}$$

Now, let $\phi(y_i|y_j)$ be the conditional distribution of $y_i|y_j$, then, for $i=1,2$, we have

$$\begin{aligned} p_i &= \int_0^t \phi(y_i-\theta_i)\int_{-\infty}^{\rho y_i}\phi(y_{3-i}|y_i)\,dy_{3-i}dy_i \\ &= \Phi\left(\frac{\rho\theta_i-\theta_{3-i}}{\sqrt{1-\rho^2}}\right)(\Phi(t-\theta_i)-\Phi(-\theta_i)) \ . \end{aligned}$$

Moreover, fixing for example $\rho \geq 0$, we have

$$\begin{aligned} p_3 &= \int_0^{\rho h}\phi(y_1-\theta_1)\,dy_1\int_{\rho y_1}^{y_1/\rho}\phi(y_2|y_1)\,dy_2 \\ &+ \int_{\rho h}^{t}\phi(y_1-\theta_1)\,dy_1\int_{\rho y_1}^{\rho y_1+\sqrt{(1-\rho^2)(t^2-y_1^2)}}\phi(y_2|y_1)\,dy_2 \ . \end{aligned}$$

This gives

$$\begin{aligned} p_3 =& M\left(\frac{\theta_1-\rho\theta_2}{\rho\sqrt{1-\rho^2}},\frac{\sqrt{1-\rho^2}}{\rho}\right)_{-\theta_1}^{\rho t-\theta_1} - \Phi\left(\frac{\rho\theta_1-\theta_2}{\sqrt{1-\rho^2}}\right)(\Phi(t-\theta_1)-\Phi(-\theta_1)) \\ &+ N\left(\frac{\rho\theta_1-\theta_2}{\sqrt{1-\rho^2}},\theta_1,t\right)_{\rho t-\theta_1}^{t-\theta_1} - \Phi\left(\frac{\rho\theta_1-\theta_2}{\sqrt{1-\rho^2}}\right)(\Phi(t-\theta_1)-\Phi(\rho t-\theta_1)) \ . \end{aligned}$$

Simplifying p_3 and p_1 we get (5).

Equation (6) follows using direction invariance in p_3 and Johnson and Kotz (1976), page 95.

6.2 Proof of Proposition 3

From (4) and

$$I_{(3)} = 1 - \sum_{s=0,1,2} I_{(s)}$$

we have

$$\begin{aligned} Q &= y_1^2 I_{(1)} + y_2^2 I_{(2)} + y'\Sigma^{-1}y - y'\Sigma^{-1}y\left(I_{(0)}+I_{(1)}+I_{(2)}\right) \\ &= y'\Sigma^{-1}y - \frac{1}{1-\rho^2}\sum_{s=0}^{2}\left(\sum_{i=1,2} y_i^2 I_{(s)}\rho^{2I(i=s)} - 2\rho y_1 y_2 I_{(s)}\right) . \end{aligned}$$

Now, using the noncentral χ^2 expectation, i.e.

$$E\left(y'\Sigma^{-1}y\right) = 2 + \theta'\Sigma^{-1}\theta$$

and the conditional gaussian density of $y_{3-i}|y_i$ for $i=1,2$, we get

$$
\begin{aligned}
&E\left(y_i^m I_{(i)}|\theta\right)\\
&= \int_0^{+\infty} y_i^m \phi\left(y_i-\theta_i\right) dy_i \int_{-\infty}^{\rho y_i} \phi\left(\frac{y_{3-i}-\theta_{3-i}-\rho\left(y_i-\theta_i\right)}{\sqrt{1-\rho^2}}\right) \frac{dy_{3-i}}{\sqrt{1-\rho^2}}\\
&= \Phi\left(\frac{\rho\theta_i-\theta_{3-i}}{\sqrt{1-\rho^2}}\right) \int_0^{+\infty} y_i^m \phi\left(y_i-\theta_i\right) dy_i\\
&= \Phi\left(\frac{\rho\theta_i-\theta_{3-i}}{\sqrt{1-\rho^2}}\right) \int_{-\theta_i}^{+\infty} \left(t+\theta_i\right)^m \phi\left(t\right) dt\\
&= \Phi\left(\frac{\rho\theta_i-\theta_{3-i}}{\sqrt{1-\rho^2}}\right) \sum_{j=0}^{m} \binom{m}{j} \theta_i^{m-j} K_{j,-\theta_i}^{+\infty}.
\end{aligned}
$$

The other expectations in proposition 3 are similarly obtained.

6.3 Proof of Proposition 4

Proof of (11). The form of $\hat{\theta}$ follows immediately observing that the sum is simply a choice operator and $\pi_s^{-1}\left(\hat{\eta}_{(s)}\right) I_{(s)}$ is the required projection expressed in the original arrangement of the coordinates as θ, if $I_{(s)}\left(y\right)=1$.

Proof of (12). To get the score (1) note that, using the fact $y-\pi^{-1}\left(z\right)=\pi^{-1}\left(\pi\left(y\right)-z\right)$, we can write

$$
\begin{aligned}
y-\hat{\theta} &= \sum_{s=1}^{p} \pi_s^{-1}\left(\begin{pmatrix} a_{(s)} \\ b_{(s)} \end{pmatrix} - \begin{pmatrix} 0_{q_s} \\ b_{(s)}-\Delta_{(s)ba}\Delta_{(s)aa}^{-1}a_{(s)} \end{pmatrix}\right) I_{(s)}\\
&= \sum_{s=1}^{p} \pi_s^{-1} \begin{pmatrix} a_{(s)} \\ \Delta_{(s)ba}\Delta_{(s)aa}^{-1}a_{(s)} \end{pmatrix} I_{(s)}\ .
\end{aligned}
$$

Recalling that $I_{(s)}I_{(r)}=0$ with probability one for $r\neq s$, it follows that the score (1) can be written as in (12) where

$$
Q_{(p)}=y'\Sigma^{-1}y
$$

and, for $s=1,...,p-1$,

$$
Q_{(s)}=\pi_{(s)}^{-1}\left(a'_{(s)},a'_{(s)}\Delta_{(s)aa}^{-1}\Delta_{(s)ab}\right)\Sigma^{-1}\pi_{(s)}^{-1}\begin{pmatrix} a_{(s)} \\ \Delta_{(s)ba}\Delta_{(s)aa}^{-1}a_{(s)} \end{pmatrix}
$$

reduces to

$$
\begin{aligned}
Q_{(s)} &= a'\left(\Sigma^{ab}\Delta_{ba}\Delta_{aa}^{-1}+\Delta_{aa}^{-1}\Delta_{ab}\Sigma^{ba}+\Delta_{aa}^{-1}\Delta_{ab}\Sigma^{bb}\Delta_{ba}\Delta_{aa}^{-1}+\Sigma^{aa}\right)a \qquad (13)\\
&= a'_{(s)}\Delta_{(s)aa}^{-1}a_{(s)}.
\end{aligned}
$$

Note that the deponent (s) has been omitted from the first line of (13) for simplicity and

$$
\pi_{(s)}\left(\Sigma^{-1}\right)=\begin{pmatrix} \Sigma^{aa} & \Sigma^{ab} \\ \Sigma^{ba} & \Sigma^{bb} \end{pmatrix}
$$

has been partitioned conformably to the other matrices.

6.4 Proof of Proposition 5

From (12) and (13) we have

$$F_Q\left(t^2\right) = P\left(I_{(0)} = 1\right) + \sum_{s=1}^{p} P\left(Q_{(s)} < t^2, I_{(s)} = 1\right) \tag{14}$$
$$= \Phi_\Sigma\left(0_k\right) + \sum_{s=1}^{p-1} P\left(a'_{(s)}\Delta^{-1}_{(s)aa}a_{(s)} < t^2, I_{(s)} = 1\right)$$
$$+ P\left(y'\Sigma^{-1}y < t^2, I_{(p)} = 1\right) \ .$$

Omitting the deponent (s) from the following relations, recalling (10) , and using multiple integral notation we have, say,

$$p_{(s)} = P\left(a'\Delta^{-1}_{aa}a < t^2, I_{(s)} = 1\right) = \int_{a'\Delta^{-1}_{aa}a<t^2,a>0} \phi_{\Delta_{aa}}\left(a\right) da \int_{-\infty}^{\Delta_{ba}\Delta^{-1}_{aa}a} \phi\left(b|a\right) db.$$

Now, letting

$$w = \Delta^{-1/2}_{aa}a$$

we have

$$p_{(s)} = \int_{w'w<t^2,\Delta^{1/2}_{aa}w>0} \phi\left(w\right) dw \int_{-\infty}^{\Delta_{ba}\Delta^{-1/2}_{aa}w} \phi\left(b|w\right) dw$$
$$= \Phi\left(0\right)_{\Delta(w)} \int_{w'w<t^2,\Delta^{1/2}_{aa}w>0} \phi\left(w\right) dw \ ,$$

using direction invariance of $w'w$ in the above RHS integral we have

$$\int_{w'w<t^2,\Delta^{1/2}_{aa}w>0} \phi\left(w\right) dw = P\left(\chi^2_{q_s} < t^2\right) \frac{Vol\left(A_{(s)}\left(t\right)\right)}{Vol\left(S_{q_s}\left(t\right)\right)}.$$

Similarly, we can handle the last term of (14) .

References

M. Basseville and I.V. Nikiforov. *Detection of Abrupt Changes*. Prentice Hall, Englewood Cliffs, 2nd edition, 1993.

D. Conniffe. Testing hypotheses with estimated scores. *Biometrika*, 77(1):97–106, 1990.

A. Fassò. Localizing ruptures in block stochastic systems. *J. Ital. Statist. Soc.*, 1(2): 235–249, 1992.

A. Fassò. One-sided multivariate control charts. In A. Azzalini and F. Parpinel, editors, *Proceedings of "Meeting in Honour of Oliviero Lessi"*. Dept. of Statistics, University of Padova, Padova, 1996.

A. Fassò. On some control charts for nonlinear ruptures. *Ital. J. Appl. Statist.*, 9(1): 123–141, 1997.

N.L. Johnson and S. Kotz. *Distributions in Statistics: Continuous Multivariate Distributions*. Wiley, New York, 1976.

R.A. Johnson and M. Bagshaw. The effect of serial correlation on the performance of $cusum$ tests. *Technometrics*, 16(1):103–112, 1974.

A. Kudo. A multivariate analogue of the one-sided test. *Biometrika*, 50(3):403–418, 1963.

C.A. Lowry, W.H. Woodall, C.W. Champ, and S.E. Rigdon. A multivariate exponentially weighted moving average control chart. *Technometrics*, 34(1):46–53, 1992.

G. Masarotto and G. Capizzi. Calibration of a control chart via stochastic approximation. *In printing on proceeding of the Italian Statistical Society Congress "XXXIX Riunione Scientifica", Sorrento*, 1998.

L. Mason, C. Champ, N. Tracy, S. Wierda, and J. Young. Assessment of multivariate process control techniques. *J. Qual. Technol.*, 29(2):140–143, 1997.

P. Nuesch. On the problem of testing location in multivariate populations for restricted alternatives. *Annals Math. Statist.*, 37:113–119, 1966.

T. Robertson, F. Wright, and R. Dykstra. *Order Restricted Statistical Inference*. Wiley, New York, 1988.

G. C. Runger and S. S. Prabhu. A Markov chain model for the multivariate exponentially weighted moving averages control chart. *J. of Amer. Statist. Ass.*, 91(436):1701–1706, 1996.

A. Shapiro. Towards a unified theory of inequality constrained testing in multivariate analysis. *Int. Statist. Rev.*, 86(1):49–62, 1988.

M. Shipper, J. Den Hartog, and E. Meelis. Sequential analysis of environmental monitoring data: Optimal $sprt$s. *Environmetrics*, 8(1):29–41, 1997.

P.C. Young and K.J. Beven. Data-based mechanistic modeling and the rainfall-flow nonlinearity. *Environmetrics*, 5(3):335–363, 1994.

Author's address:

Prof. Alessandro Fassò
Department of Engeneering
University of Bergamo
Via Marconi 5
I-24044 Dalmine BG
Italy

Tel. +39 35 277 323
Fax +39 35 562 779
email: fasso@unibg.it