Research Article

On directed alternatives in linear inference

Abstract

Tests for vector hypotheses $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ in $\mathbb{R}^k$ typically have powers depending on quadratic forms of type $\lambda = (\theta - \theta_0)' \Xi^{-1} (\theta - \theta_0)$. This study examines the case that $(\mu - \mu_0)$ is restricted to subspaces, for example, $(\mu - \mu_0)' = c \, (1,1,0,\ldots,0)$ differing only in their first two coordinates. These are called directed alternatives. The spectral decomposition of   supports the identification of one–dimensional alternatives least likely and most likely to be discerned, to complement conventional data analysis. Applications are drawn in the use of Hotelling's T^2 and of F–tests in linear inference. Moreover, it is seen that a given design may be recast so as to reverse the least likely and most likely alternatives. Numerical examples serve to illustrate the findings.

$62J05, \; 62H10 \; and \; 62P30$

Keywords: linear models; F tests, hotelling's T^2 tests, directed alternatives, reversal designs

Donald Jensen
Department of Statistics, Virginia Tech, USA

Correspondence: Donald Jensen, Department of Statistics, Virginia Tech, Blacksburg, VA 24061, USA, Email djensen@vt.edu

Introduction

Power in statistical inference is driven by non–null distributions. For observations in $\mathbb{R}^k$ having dispersion matrix Ξ, noncentrality parameters often emerge as the Mahalanobis[1] distance between points (u, v) in $\mathbb{R}^k$, namely,

$$D_{\Xi}^2 (u,v) = (u-v)' \Xi^{-1} (u-v). \qquad (1)$$

This specializes to the Euclidean metric for the case that $\Xi = c \, I_k$, in which case the model is called isotropic. In particular, nonparametric and other statistics often have noncentral chi–squared distributions, either in small samples or asymptotically. In addition, pervasive venues in parametric inference, to be reexamined in some detail, include the following.

Case (i). Hotelling[2] Test: $T^2 = n (\bar{Y} - \mu_0) S^{-1} (\bar{Y} - \mu_0)$ where $(\bar{Y}, S)$ are the sample mean and dispersion matrix of n Gaussian vectors in $\mathbb{R}^k$ having the location–scale parameters   . Then in testing $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$, the power function is $\Psi(\lambda)$ with noncentrality $\lambda = n (\mu - \mu_0)' \Sigma^{-1} (\mu - \mu_0)$.

Case (ii). The General Linear Model: $\{Y = X\beta + \varepsilon\}$ with Gaussian errors having zero means and dispersion matrix $\sigma^2 I_n$. Then in testing $H_0 : \beta = \beta_o$ against $H_1 : \beta \neq \beta_o$ in $\mathbb{R}^k$, the power function is $\Psi(\lambda)$ with noncentrality $\lambda = (\beta - \beta_o)' X'X (\beta - \beta_o) / \sigma^2$.

Classical theory allows for any $(\mu - \mu_0) \in \mathbb{R}^k$ for Case (i), and $(\beta - \beta_o) \in \mathbb{R}^k$ for Case (ii). On the other hand, alternatives lying in designated subspaces may hold substantive interest per se. For example, taking $(\mu - \mu_0)' = c \, (1,1,0,\ldots,0)$ allows for discrepancies between μ and μ_0 in their first two coordinates only, whereas $(\mu - \mu_0)' = c \, (1,1,\ldots,1)$ allows for deviations along the equiangular line in $\mathbb{R}^k$. Both are one–dimensional; subspaces of dimension greater than one are considered subsequently. Alternatives lying in designated subspaces of $\mathbb{R}^k$ are called directed alternatives, and the goal here is to study powers of tests against alternatives of these types.

The present study expands on this as follows. Not only do distinct alternatives differ in importance to users, but so too their probabilities of detection. Here the spectral decomposition of Ξ, if *anisotropic*, supports the identification of alternatives least likely and most likely to be discovered, as well as intermediate cases. These serve to bracket the effective range of inferences intrinsic to a given study, and thereby complement conventional options in data analysis. Applications are drawn in the use of Hotelling's T^2 in multivariate samples, and of F–tests in the analysis of linear models. Moreover, it is shown that a given design may be modified so as to reverse the least likely and most likely alternatives, in the event that this would better serve the objectives of an experiment.

This study is organized as follows. Supporting developments are given next in Section 2, followed by the principal findings of Section 3. Several examples in Section 4 illustrate the essential results. Collateral materials are deferred for completeness to an Appendix.

Preliminaries

Notation

Spaces include $\mathbb{R}^n$ as Euclidean n-space; $\mathbb{R}_+^n$ as its positive orthant; $\mathbb{S}_n$ as the real symmetric $(n \times n)$ matrices; $\mathbb{S}_n^+$ as their positive definite varieties; $\mathbb{F}_{n \times k}$ as the real $(n \times k)$ matrices of rank $k \leq n$; and $\mathcal{O}_k$ as the $(k \times k)$ orthogonal group. Vectors and matrices are set in bold type; the transpose, inverse, trace, and determinant of A are A', A^{-1}, $tr(A)$, and $|A|$; the unit vector in $\mathbb{R}^n$ is $\mathbf{1}_n = [1,\ldots,1]'$; I_n is the $(n \times n)$ identity; and $\text{Diag}(A_1,\ldots,A_k)$ is a block-diagonal array. If $B = [b_1,\ldots,b_k]$ is of order $(n \times k)$ and rank $k < n$, then $S_p(B)$ designates the column span of B, i.e., the k–dimensional subspace of $\mathbb{R}^n$ spanned by $[b_1,\ldots,b_k]$. The ordered eigenvalues of $A \in \mathbb{S}_n$ are $\{\lambda_i(A) = \alpha_i; \; 1 \leq i \leq n\}$ with $\{\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n\}$, and its spectral decomposition is $A = PD_\alpha P' = \Sigma_{i=1}^n \alpha_i p_i p_i'$, where $P = [p_1,\ldots,p_n] \in \mathcal{O}_n$ and $D_\alpha = \text{Diag}(\alpha_1,\ldots,\alpha_n)$. By convention its condition number is $C_1(A) = \alpha_1 / \alpha_n$. The singular decomposition

of $\boldsymbol{B} \in \mathbb{F}_{n \times k}$ is $B = \boldsymbol{P}\boldsymbol{D}_{\delta}\boldsymbol{Q}' = \Sigma_{i=1}^{k} \delta_i \boldsymbol{p}_i \boldsymbol{q}_i'$, where the mutually orthogonal columns of $\boldsymbol{P} = \left[\boldsymbol{p}_1,...,\boldsymbol{p}_k\right]$ comprise the left–singular vectors; $D_{\delta} = \mathrm{Diag}\left(\delta_1,...,\delta_k\right)$ are its singular values; and columns of $\boldsymbol{Q} \in \mathcal{O}_k$ are the right–singular vectors.

Special distributions

For $\boldsymbol{Y} \in \mathbb{R}^n$, its distribution, mean, and dispersion matrix are L($\boldsymbol{Y}$), $E\left(\boldsymbol{Y}\right) = \mu$ and $V\left(\boldsymbol{Y}\right) = \Sigma$, say, with variance $Var\left(Y\right) = \sigma^2$ on $\mathbb{R}^1$. Specifically, L($\boldsymbol{Y}$) $= N_n\left(\mu, \Sigma\right)$ is Gaussian on $\mathbb{R}^n$ with parameters $\left(\mu, \Sigma\right)$. Distributions on $\mathbb{R}^1_+$ include the $\chi_2(\cdot; \nu, \lambda)$ with ν degrees of freedom and noncentrality parameter λ; the Snedecor–Fisher $F(\cdot; \nu_1, \nu_2, \lambda)$ with degrees of freedom $\left(\nu_1, \nu_2\right)$ and noncentrality λ; and Hotelling^2 $T_k^2\left(\cdot, \nu, \lambda\right)$ of order k having ν degrees of freedom and noncentrality λ. Recall that $F(\cdot; \nu_1, \nu_2, \lambda)$ increases stochastically with increasing λ with other parameters held fixed. Identify c_α in context as the upper α –level rejection rule. The power of a test, to be considered as a function of λ, is designated by $\psi(\lambda)$.

The principal findings

Directed alternatives

Our notation encompasses both (i) Hotelling^2 T^2 and (ii) General Linear Models, having location–scale parameters $\left(\delta, \Xi\right)$. What distinguishes this study are directed alternatives with examples as noted, but expanded to include alternatives $\left\{\theta_j; 1 \leq j \leq k\right\}$ aligned with the orthonormal eigenvectors $\boldsymbol{Q} = \left[\boldsymbol{q}_1,...,\boldsymbol{q}_k\right]$ of Ξ, thus standardized to unit lengths. To continue, as $\Omega = \Xi^{-1}$ assumes a central role, take $\Omega = \Sigma_{i=1}^{k} \kappa_i \boldsymbol{q}_i \boldsymbol{q}_i' = \boldsymbol{Q}\boldsymbol{D}_\kappa \boldsymbol{Q}' \in \mathbb{S}_k^+$ as its spectral decomposition, with $\left\{\kappa_1 \geq ... \geq \kappa_k > 0\right\}$. As in Appendix A.1, undertake the expansions

$$\Omega = \boldsymbol{Q}_1 \boldsymbol{D}_{\kappa_1} \boldsymbol{Q}_1' + \boldsymbol{Q}_2 \boldsymbol{D}_{\kappa_2} \boldsymbol{Q}_2' + \boldsymbol{Q}_3 \boldsymbol{D}_{\kappa_3} \boldsymbol{Q}_3' \qquad (2)$$

$$\boldsymbol{\Omega} = \boldsymbol{Q}_1 \boldsymbol{D}_{\kappa_1} \boldsymbol{Q}_1' + k_r \boldsymbol{Q}_2 \boldsymbol{D}_{\kappa_2} \boldsymbol{Q}_2' + \boldsymbol{Q}_3 \boldsymbol{D}_{\kappa_3} \boldsymbol{Q}_3'; k_r \text{ repeated } s \text{ times}, \qquad (3)$$

where elements of $\boldsymbol{Q} = [\boldsymbol{Q}_1, \boldsymbol{Q}_2, \boldsymbol{Q}_3]$ are of orders $\left\{(k \times (r-1)), (k \times s), (k \times d)\right\}$ with $d = k - r - s + 1$, and where $\boldsymbol{D}_\kappa = \mathrm{Diag}\left(\boldsymbol{D}_{\kappa_1}, \boldsymbol{D}_{\kappa_2}, \boldsymbol{D}_{\kappa_3}\right)$ is partitioned conformably. In regard to quadratic forms of type $Q(u) = u'\Omega u$ serving as noncentrality parameters, a principal result is the following.

Theorem 1. Given is a location–scale model with parameters $\left(\delta, \Xi\right)$, together with a test for $H_0: \delta = \delta_0$ against $H_1: \delta \neq \delta_0$ having power $\Psi(\lambda)$ increasing monotonically with $\lambda = D_\Xi^2\left(\delta, \delta_0\right) = \left(\delta - \delta_0\right)' \Xi^{-1}\left(\delta - \delta_0\right)$. Take $\left\{\left(\delta - \delta_0\right) = \theta_j; 1 \leq j \leq k\right\}$ in

succession as the eigenvectors $\left\{\theta_j = \boldsymbol{q}_j; 1 \leq j \leq k\right\}$ of Ξ^{-1} with eigen values $\left\{\kappa_1 \geq ... \geq \kappa_k > 0\right\}$.

i. Then powers $\Psi\left(\lambda_j\right)$ of the test at alternatives $\left\{(\delta - \delta_0) = \theta_j; 1 \leq j \leq k\right\}$ depend on the noncentrality parameters $\left\{\lambda_j = \kappa_j; 1 \leq j \leq k\right\}$, respectively.

ii. In particular, the alternatives most likely and least likely to be discerned in terms of power are θ_1 and θ_k having powers $\Psi\left(\kappa_1\right)$ and $\Psi\left(\kappa_k\right)$, respectively.

iii. Consider alternatives $\{\gamma_j \in S_p\left(\boldsymbol{Q}_2\right)\}$ standardized to unit lengths. Then bounds on powers at these local alternatives are given by $\Psi\left(\kappa_{r+s}\right) \leq \{\Psi\left(\lambda_j\right) \text{ for every } \gamma_j \in S_p\left(\boldsymbol{Q}_2\right)\} \leq \Psi\left(\kappa_r\right)$

iv. Suppose that k_r is repeated s times as in the spectral resolution (2) for Ω. Then for each alternative $\{\gamma_j = \boldsymbol{Q}_2 \boldsymbol{a}_j \in S_p\left(\boldsymbol{Q}_2\right), \text{ with } \boldsymbol{a}_j' = \left[a_{j_1},...,a_{j_s}\right]\}$, the noncentrality parameter is $\left\{\lambda_j = \kappa_r \boldsymbol{a}_j' \boldsymbol{a}_j\right\}$, with corresponding power $\left\{\Psi\left(\lambda_j\right): \gamma_j \in S_p\left(\boldsymbol{Q}_2\right)\right\}$.

Proof: Conclusion (i) follows directly from $\lambda\left(\theta_j\right) = \theta_j' \Xi^{-1} \theta_j = \theta_j'\left(\Sigma_{i=1}^k \kappa_i \boldsymbol{q}_i \boldsymbol{q}_i'\right)\theta_j = \kappa_j$ since $\theta_j = \boldsymbol{q}_j$, whereas $\left\{\boldsymbol{q}_j' \boldsymbol{q}_i = 0; i \neq j\right\}$ and $\boldsymbol{q}_j' \boldsymbol{q}_j = 1$ by orthonormality. Conclusion (ii) follows directly from variational properties of Rayleigh quotients as in Lemma A.1(i) of the Appendix. In like manner conclusion (iii) follows from Lemma A.1(ii) as variational properties over subspaces. Conclusion (iv) follows from (iii) since $\left\{\boldsymbol{q}_j' \boldsymbol{q}_i = 0; i \neq j\right\}$

Remark 1. The directed alternatives $\delta_1' = [1,1,0,...,0]$ and $\delta_2' = [1,1,...,1]$ were featured earlier as discrepancies in the first two coordinates of $\left(\delta - \delta_0\right)$, and as deviations about the equiangular line in $\mathbb{R}^k$. Let $\Xi^{-1} = \left[\xi^{ij}\right]$. Then powers $\Psi(\lambda)$ at these alternatives will depend on $\lambda_1 = \sum_{i=1}^2 \sum_{j=1}^2 \xi^{ij}$ at $\delta_1' = [1,1,0,...,0]$ and on $\lambda_2 = \sum_{i=1}^k \sum_{j=1}^k \xi^{ij}$ at $\delta_2' = [1,1,...,1]$.

Corollary 1. On specializing the location–scale parameters $\left(\delta, \Xi\right)$, Theorem 1 applies verbatim as follows.

(i) Hotelling^2 T^2: $\left(\delta, \Xi\right) = \left(\mu, \Sigma\right); L\left(T^2\right) = T_k^2\left(\cdot; n-1, \lambda\right)$, the power $\Psi(\lambda)$ depending on $\lambda = n\left(\mu - \mu_0\right)' \Sigma^{-1}\left(\mu - \mu_0\right)$.

(ii) General Linear Models: $\left(\delta, \Xi\right) = \left(\beta, \left(\boldsymbol{X'X}\right)^{-1}\right); L(F) = F\left(\cdot; k, n-k, \lambda\right)$, the power $\Psi(\lambda)$ depending on $\lambda = \left(\beta - \beta_o\right)' \boldsymbol{X'X}\left(\beta - \beta_o\right)/\sigma^2$.

Proof: The noncentral distribution $F\left(\cdot; k, n-k, \lambda\right)$ clearly satisfies the assumptions of Theorem 1 on identifying $\left(\delta, \Xi\right) = \left(\beta, \left(\boldsymbol{X'X}\right)^{-1}\right)$ as claimed. Similarly in testing $H_0: \mu = \mu_0$ against $H_1: \mu \neq \mu_0$, Hotelling's T^2 inherits these properties through the conversion

$$L\left(\frac{(\nu-k+1)T^2}{k\nu}\right)=F(;k,\nu-k+1,\lambda). \text{ With } \lambda=n(\boldsymbol{\mu}-\boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}-\boldsymbol{\mu}_0).$$

Remark 2. Note that alternatives $\{\boldsymbol{\theta}_j=\boldsymbol{q}_j;1\le j\le k\}$ of unit lengths give noncentrality parameters $\{\kappa_j;1\le j\le k\}$. If instead the directed alternatives are $\{\boldsymbol{\theta}_j=c_j\boldsymbol{q}_j;1\le j\le k\}$, then the noncentrality parameters will be $\{c_j^2\kappa_j;1\le j\le k\}$.

Remark 3. Note that the foregoing developments are for the general case that $\Omega=\boldsymbol{QD}_\kappa\boldsymbol{Q}'$ is anisotropic with $\{\kappa_1\ge\ldots\ge\kappa_k>0\}$. If isotropic, then the following applies.

Definition 1. The model $\boldsymbol{\Xi}$ is called isotropic if and only if $\boldsymbol{\Xi}=d\boldsymbol{I}_k$, in which case power functions are directionally invariant, not depending on directions of alternatives in $\mathbb{R}^k$. Add: Bounds on ARLs from restricted variation.

Sphericity

The density for $N_k(\boldsymbol{\mu},\boldsymbol{\Sigma})$ has spherical contours for the case that $\boldsymbol{\Sigma}=d\boldsymbol{I}_k$, i.e., the model is isotropic. Sample evidence regarding the isotropy of T^2 is available. Mauchly[3] derived the Likelihood Ratio test for sphericity, namely, $H_0:\boldsymbol{\Sigma}=d\boldsymbol{I}_k$ against $H_1:\boldsymbol{\Sigma}\ne d\boldsymbol{I}_k$. A contemporary test utilizes the modified statistic

$$LR_M=-\left(\nu-\frac{2k^2+k+2}{6k}\right)ln\left[\frac{k^k|\boldsymbol{S}|}{(tr\boldsymbol{S})^k}\right] \qquad (4)$$

taking $\boldsymbol{S}$ as the sample dispersion matrix from n observations, rejecting at level α for $LR_M>c_\alpha$ with $\nu=n-1$ and with c_α as the upper percentile of the central distribution $\chi^2(k(k+1)/2-1,0)$. See, for example, Rencher.[4]

Design reversals

Developments thus far are predicated in part on the desirability to identify alternatives having varying powers of discernment. These include the most likely and least likely as in Theorem 1(ii). If the least likely is deemed to be of greatest interest, it remains to ask whether it might serve instead as the most likely alternative. In the context of designed experiments the answer is affirmative, as the intrinsic structure offers a venue for modifying a given design so as to achieve these ends. Details follow.

Consider the model $\{\boldsymbol{Y}=[\boldsymbol{1}_n,\boldsymbol{X}][\alpha,\beta']'+\boldsymbol{\varepsilon}\}$ with $\boldsymbol{X}$ centered such that $\boldsymbol{1}_n'\boldsymbol{X}=0$, where location–scale parameters for $\hat{\boldsymbol{\beta}}$ are $(\beta,\boldsymbol{\Xi})$ with $\boldsymbol{\Xi}=(\boldsymbol{X}'\boldsymbol{X})^{-1}$ as in Corollary 1(ii). In particular, the test for $H_0:\beta=\beta_o$ against $H_1:\beta\ne\beta_o$ utilizes $F=(\hat{\boldsymbol{\beta}}-\beta_o)'\boldsymbol{X}'\boldsymbol{X}(\hat{\boldsymbol{\beta}}-\beta_o)/kS^2$ with S^2 as the residual mean square and with noncentrality $\lambda=(\beta-\beta_o)'\boldsymbol{X}'\boldsymbol{X}(\beta-\beta_o)/\sigma^2$, where it often suffices to take $\sigma^2=1.0$. For $\boldsymbol{X}\in\boldsymbol{F}_{n\times k}$ its singular decomposition, followed by $\boldsymbol{X}'\boldsymbol{X}$, is

$$\boldsymbol{X}=\boldsymbol{PD}_\delta\boldsymbol{Q}'=\sum_{i=1}^k\delta_i\boldsymbol{p}_i\boldsymbol{q}_i',\boldsymbol{X}'\boldsymbol{X}=\boldsymbol{QD}_\kappa\boldsymbol{Q}' \qquad (5)$$

with $\boldsymbol{P}=[p_1,\ldots,p_k]$ as its left–singular vectors, $\boldsymbol{D}_\delta$ as its singular values, and columns of $\boldsymbol{Q}\in\mathcal{O}_k$ as its right–singular vectors. Clearly $\{\kappa_i=\delta_i^2;1\le i\le k\}$. Our principal reconstruction is articulated in the following.

Theorem 2. Let $\pi(\boldsymbol{\delta})$ be a permutation operator reversing the ordered array $[\delta_1\ge\ldots\ge\delta_k]$ to $\{\delta_k\le\{\delta_2,\ldots,\delta_{k-2}\}\le\delta_1\}$, and let $\boldsymbol{D}_\pi=\text{Diag}(\delta_k,\delta_2,\ldots,\delta_{k-2},\delta_1)$. Next construct

$$\boldsymbol{X}_\pi=\boldsymbol{PD}_\pi\boldsymbol{Q}'=\delta_k\boldsymbol{p}_1\boldsymbol{q}_1'+\sum_{i=2}^{k-1}\delta_i\boldsymbol{p}_i\boldsymbol{q}_i'+\delta_1\boldsymbol{p}_k\boldsymbol{q}_k'$$

such that pairs $(\delta_k,\boldsymbol{q}_1)$ and $(\delta_1,\boldsymbol{q}_k)$ are realigned.

Conclusion: The most likely and least likely alternatives for design $\boldsymbol{X}_\pi$ are reversed from those of $\boldsymbol{X}$, so that $\boldsymbol{\theta}_1=\boldsymbol{q}_k$ now is most likely with power $\Psi(\kappa_1)$ depending on κ_1, and $\boldsymbol{\theta}_k=\boldsymbol{q}_1$ least likely with power $\Psi(\kappa_k)$ depending on κ_k.

Proof: Clearly the conventional reordering of eigenvalues gives

$$\boldsymbol{X}_\pi'\boldsymbol{X}_\pi=\kappa_1\boldsymbol{q}_k\boldsymbol{q}_k'+\sum_{i=2}^{k-1}\kappa_i\boldsymbol{q}_i\boldsymbol{q}_i'+\kappa_k\boldsymbol{q}_1\boldsymbol{q}_1' \qquad (6)$$

and the conclusion follows on applying Theorem 1(ii) in the context of Corollary 1(ii).

Remark 4. Variations on $\boldsymbol{X}_\pi$ are apparent. Any permutation of $\{\delta_2,\ldots,\delta_{k-1}\}$ gives the same conclusion. In addition, any pairs $\{(\delta_i,\boldsymbol{q}_i),(\delta_j,\boldsymbol{q}_j)\}$ may be selected in like manner as most likely and least likely to be discerned. Note, however, that these tools are available in the case of first–order designs.

Case studies

Studies in Hotelling[2] T^2 and second–order response models are given, to illustrate Theorem 1 and Corollary 1. Moreover, an example design serves to illustrates the Theorem 2 reversal of most likely and least likely alternatives.

Hotelling's T^2 tests

We reexamine the role of calcium in the growth of turnip greens, using data as reported in Kramer et al.,[5] In each of 29 experimental plots the plant calcium (Y_1) was determined, and the soil calcium was assayed as available (Y_2) and exchangeable (Y_3) calcium. The units all are milliequivalents per hundred grams. Horticultural specialists expect these to run at about 15.00, 6.00 and 2.85 units, respectively. The sample means are $\bar{\boldsymbol{Y}}'=[17.97,4.39,2.46]$; and the sample dispersion matrix is $\boldsymbol{S}$ with inverse $\boldsymbol{S}^{(-1)}=\boldsymbol{QD}_\kappa\boldsymbol{Q}'$ in spectral form, as listed in

$$Q=\begin{bmatrix} 0.01024 & 0.22479 & -0.97435 \\ 0.05973 & -0.97280 & -0.22381 \\ -0.99816 & -0.05591 & -0.02339 \end{bmatrix}$$

where $\boldsymbol{D}_\kappa = Diag$ (5.61102, 0.04857, 0.00416) The data are ill–conditioned, with condition number $c_1(\boldsymbol{S}) = 1,348.80$,

The statistic reported is $T^2 = n(\overline{\boldsymbol{Y}} - \boldsymbol{\mu}_0)'\boldsymbol{S}^{-1}(\overline{\boldsymbol{Y}} - \boldsymbol{\mu}_0) = 24.97$ with $\boldsymbol{\mu}_0' = [15.00, 6.00, 2.85]$, rejecting at level $\alpha = 0.01$ the hypothesis $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$ in favor of some $H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$. Indeed, the p–value is $P(T^2 > 24.97 \mid H_0) = 0.000751$ with $C_\alpha = [14.980]$

On applying Corollary 1(i) with $\boldsymbol{S}$ in lieu of $\boldsymbol{\Sigma}$, the columns of $\boldsymbol{Q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3]$ are taken as successive alternatives to $(\boldsymbol{\mu} - \boldsymbol{\mu}_0)$, namely

$$\begin{bmatrix} \mu_1 - 15.00 \\ \mu_2 - 6.00 \\ \mu_3 - 2.85 \end{bmatrix} \in \begin{bmatrix} 0.01024 & 0.22479 & \mathbf{-0.97435} \\ 0.05973 & \mathbf{-0.97280} & -0.22381 \\ \mathbf{-0.99816} & -0.05591 & -0.02339 \end{bmatrix}$$

where the dominant terms are in bold type. The noncentrality parameters $\{\lambda_i = n\kappa_i; 1 \leq i \leq 3\}$ are $\{162.720, 1.409, 0.121\}$, and taking $\alpha = 0.05$ and $C_\alpha = 9.612$, , powers at these alternatives are

$P((T^2 \mid \lambda_1) > 9.612 = 1.0000, P((T^2 \mid \lambda_2) > 9.612 = 0.1316, P((T^2 \mid \lambda_3) > 9.612) = 0.0562.$

Accordingly, T^2 has essentially unit power to distinguish the hypothetical deviation $(\mu_3 - 2.85)$ from -0.99816, since the discrepancies 0.01024 for $(\mu_1 - 15.00)$ and 0.05973 for $(\mu_2 - 6.00)$ in q_1 are negligible. Similarly, T^2 is marginally able to distinguish $[(\mu_1 - 15.00), (\mu_2 - 6.00)]$ from $[0.22479, -0.97280]$ with power 0.1316, but is virtually unable to separate $[(\mu_1 - 15.00), (\mu_2 - 6.00)]$ from $[-0.97435, -0.22381]$ with negligible power of 0.0562. In short, the latter suggests [14.03, 5.78] to be plausible values for $[\mu_1, \mu_2]$.

This is an example, as seen subsequently also, where elements of $\boldsymbol{Q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3]$, especially $\boldsymbol{q}_1$, convey useful information in regard to the objectives of the study. In summary, details regarding directed alternatives, enabled here by Theorem 1 and Corollary 1(i), go beyond conventional useage for T^2.

Hotelling's T^2 charts.

Multivariate diagnostics figure prominently in Statistical Process Control (SPC), as reviewed subsequently. In monitoring the manufacture of bomb sights during World War II, Hotelling[6] devised T^2 charts for multivariate means in $\mathbb{R}^k$ Here successive values $\{T^2; i = 1, 2, ...\}$ are charted against time, where the chart signals the process to be out–of–control at level α whenever $T_i^2 > C_\alpha$.

Moreover, with power $\psi(\lambda) = P((T_i^2 \mid \lambda) > C_\alpha)$, the Average Run Length (ARL) of time–to–signal is $ARL = 1/\psi(\lambda)$. To monitor the mean $\boldsymbol{\mu}$ against its target value $\boldsymbol{\mu}_0$, successive samples of size n yield $(\overline{\boldsymbol{Y}_i}, \boldsymbol{S}_i)$, together with $T_i^2 = n(\overline{\boldsymbol{Y}_i} - \boldsymbol{\mu}_0)'\boldsymbol{S}_i^{-1}(\overline{\boldsymbol{Y}_i} - \boldsymbol{\mu}_0)$ having the $T_k^2(\therefore n-1, \lambda)$ distribution with $\lambda = n(\boldsymbol{\mu} - \boldsymbol{\mu}_0)'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - \boldsymbol{\mu}_0)$. Phase I in SPC is set to establish base line process capabilities, to include parameter estimation, followed in Phase II by implementing the control charts themselves.

To continue, consider the data of Quesenberry[7] to be in Phase I, comprising n=30 records of 11 quality characteristics indexed in time–order of production. Following Williams et al.,[8] dimensions are reduced on selecting the first k=5 quality characteristics, namely, $[Y_1, ..., Y_5]$ having means $[\mu_1, ..., \mu_5]$, respectively. As in Section *Hotelling's T^2 Tests* we take $\boldsymbol{S}$ in lieu of $\boldsymbol{\Sigma}$, finding the spectral resolution $\boldsymbol{S}^{-1} = \boldsymbol{QD}_\kappa\boldsymbol{Q}'$ as reported in Table 1. The data are seen to be highly ill–conditioned, with condition number $c_1(\boldsymbol{S}) = 3,945.64$. In keeping with Corollary 1(i), five directed alternatives comprise the columns of $\boldsymbol{Q} = [\boldsymbol{q}_1, ..., \boldsymbol{q}_5]$ in Table 1, where dominant elements again are in bold type. In particular, this example shows $\{\boldsymbol{q}_1, ..., \boldsymbol{q}_5\}$ to be separately informative *per se*, as each corresponds essentially to deviations in $(\mu_i - \mu_{i0})$ for observations $\{Y_1, Y_3, Y_5, Y_2, Y_4\}$, respectively, since values other than those in bold type are negligible.

Taking $\{\lambda_i = n\kappa_i; 1 \leq i \leq 5\}$ with $n = 30$, $\alpha = 0.05$, and critical value $C_\alpha = 15.097$, powers against these directed alternatives are listed in Table 1. These show that deviations in the directions of $\{\mu_1, \mu_2, \mu_3, \mu_5\}$ essentially would be discerned with power at least 0.9882, whereas power in the direction of μ_4 would be diminished to 0.2367. Note, however, that these values are inflated by the value n=30 in Phase I. Samples much smaller in size ordinarily would be taken in Phase II, say n=8 in this example. Then the powers corresponding to $\{\lambda_i = 8k_i; 1 \leq i \leq 5\}$ at α=0.05 are given subsequently in Table 1. In short, with $ARL = 1/\psi(\lambda)$ as in the final row of Table 1 with samples of size n=8, these charts would detect changes in μ_1 and μ_3 immediately on average, but less responsive otherwise with $ARLs of (5.41, 5.62, 15.50) for (\mu_5, \mu_2, \mu_4)$, respectively.

Remark 5. This example goes beyond conventional uses of T^2 charts, in demonstrating that the capacity of a given chart to detect alternatives may differ widely across alternatives in $\mathbb{R}^5$ of compelling practical interest. In short, each of five ARLs pertains here to an informative one–dimensional alternative.

Second–order designs

Second–order models of type

$$\{Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_{11} X_{1i}^2 + \beta_{22} X_{2i}^2 + \beta_{12} X_{1i} X_{2i} + \varepsilon_i; 1 \leq i \leq n\} \tag{8}$$

are considered having zero mean, uncorrelated errors with variance σ^2 In a typical setting the yield (Y_i) of a chemical process is examined at specified reaction time (X_{1i}) and temperature (X_{2i}). Small designs

of historical consequence are the Central Composite (CCD) designs of Box et al.,[9] having design points as listed in Table 2.

Table 1 Spectral values for $S^{-1} = QD_\kappa Q'$ for the data of Quesenberry (2001) of order (30×5)

Eigenvalues				
κ_1	κ_2	κ_3	κ_4	κ_5
520.4304	11.9529	1.1404	1.0843	0.1319
Eigenvectors				
q_1	q_2	q_3	q_4	q_5
0.99893	−0.04553	0.00490	0.00543	0.00291
−0.00607	−0.00445	0.14300	**0.98712**	−0.07126
0.04535	**0.99863**	0.01937	0.00322	0.01722
−0.00396	−0.01681	−0.02275	0.07515	**0.99676**
−0.00505	−0.01907	**0.98926**	−0.14105	0.03287
Power at α =0.05 and n =30				
1.0000	1.0000	0.9914	0.9882	0.2367
Power at α =0.05 and n =8				
1.0000	1.0000	0.1848	0.1779	0.0645
ARLs at α =0.05 and n =8				
1.0000	1.0000	5.41	5.62	15.50

Table 2 Regressor vectors for the *CCD* design X' of order (2×9), where $\alpha = \sqrt{2}$

X'	-1.00	-1.00	$-\alpha$	α	0.00	0.00	1.00	1.00	0.00
	-1.00	1.00	0.00	0.00	$-\alpha$	α	-1.00	1.00	0.00

Proceeding as in Corollary 1(ii), we seek spectral values for the Fisher Information Matrix $X'X$, specifically, the eigenvalues and eigenvectors as listed in Table 3, with dominant terms again in bold type. Powers at these directed alternatives follow on taking $\{(\beta - \beta_o) = q_i; 1 \le i \le 6\}$ as surrogates in the noncentrality parameters for $F = (\hat{\beta} - \beta_o)' X'X (\hat{\beta} - \beta_o) / kS^2$ with $k = 6$ and S^2 as the residual mean square. Owing to only two degrees of freedom for error, computations replaced S_i^2 by $\sigma^2 = 1.0$, then powers were determined from scaled noncentral $\chi^2(6, \kappa_i)$ distributions with noncentrality parameters as listed in Table 3.

Arranged in decreasing order of their powers, these are q_1 with power 0.9765; $\{q_2, q_3, q_4\}$ each with power 0.5307; and q_5 and q_6 as alternatives with powers 0.2698 and 0.0775. Here elements of Q are separately informative: q_2 for discrepancies between (β_{11}, β_{22}) and their hypothetical values; and (q_3, q_4, q_5) for discrepancies between $(\beta_1, \beta_2, \beta_{12})$ and their hypothetical values, respectively. To continue,

observe that the eigenvalue $\kappa_2 = 8.0000$ is repeated three times. On applying Theorem 1(iv) in the context of Corollary 1(ii), we see that all standardized elements in $S_p(q_2, q_3, q_4)$ have power 0.5307. These include, in addition to q_3 for β_1 and q_4 for β_2, the standardized sums

$$\theta_1 = (q_3 + q_4) / \sqrt{2} = [0, 1, -1, 0, 0, 0]' / \sqrt{2} \qquad (9)$$

$$\theta_2 = (q_2 + q_3 + q_4) / \sqrt{3} = [0, 1, -1, 0, 7071, -0.7071, 0]' / \sqrt{3} \qquad (10)$$

for example, with θ_1 for the discrepancy between $[(\beta_1 - \beta_{10}), (\beta_2 - \beta_{20})]$ and $[1, -1] / \sqrt{2}$.

In summary, the directed second–order alternatives treated here are innovations not found in classical linear inference. Instead, these are enabled by Theorem 1 and Corollary 1(ii). Again the elements of Q, especially $\{q_2, ..., q_5\}$, are separately informative about coefficients of the model (8). Their simple and revealing structure may be attributed to the symmetry and balance of CCD designs.

Design reversals

Begin with $\{Y = [\mathbf{1}_n, X][\beta_0, \beta']' + \varepsilon\}$ with $X_0 = [\mathbf{1}_n, X]$ and X in centered form as in Section 3.2, with $\beta' = [\beta_1, \beta_2, \beta_3]$, having the design $X = P\mathrm{Diag}(\delta_1, \delta_2, \delta_3)Q'$ as listed in Table 4. Construct $X_1 = P\mathrm{Diag}(\delta_3, \delta_2, \delta_1)Q'$ on permuting singular values, but retaining the left and right singular vectors.

To continue, take the right–singular vectors, now $Q = [\theta_1, \theta_2, \theta_3]$, as directed alternatives to $(\beta - \beta_o)' = [(\beta_1 - \beta_{10}), (\beta_2 - \beta_{20}), (\beta_3 - \beta_{30})]$, together with $\kappa' = [14.5909, 4.4066, 0.2828]$ as $\{\kappa_i = \delta_i^2; 1 \le i \le 3\}$. Here the level 0.05 critical value is 6.5914; then powers are determined in turn from $F(\therefore 3, 4, \kappa_i)$ together with the power functions $\{\psi(\lambda) = P(F > 6.5914 | \kappa_i); 1 \le i \le 3\}$ having values listed in the final row of Table 4.

Table 3 Spectral values for the Fisher Information Matrix $X'X$ for the CCD design

Eigenvalues					
κ_1	κ_2	κ_3	κ_4	κ_5	κ_6
24.3427	8.0000	8.0000	8.0000	4.0000	0.6573
Eigenvectors					
q_1	q_2	q_3	q_4	q_5	q_6
0.59349	0.00000	0.00000	0.00000	0.00000	**0.80484**
0.00000	0.00000	**1.00000**	0.00000	0.00000	0.00000
0.00000	0.00000	0.00000	**-1.00000**	0.00000	0.00000
0.56911	**0.70711**	0.00000	0.00000	0.00000	**-0.41966**
0.56911	**-0.70711**	0.00000	0.00000	0.00000	**-0.41966**
0.00000	0.00000	0.00000	0.00000	**1.00000**	0.00000
Power at α =0.05					
0.9765	0.5307	0.5307	0.5307	0.2698	0.0775

Table 4 Design matrix $X = PD_\delta Q'$ and the modified $X_1 = [P\mathrm{Diag}(\delta_3, \delta_2, \delta_1)Q']$ the left (P') and right (Q) singular vectors; and the singular values $\delta' = [3.8198, 2.0992, 0.5318]$.

Design								
	-1.0000	1.0000	-1.4142	1.4142	-1.0000	1.0000	0.0000	0.0000
X'	0.0000	0.0000	1.0000	1.0000	0.0000	0.0000	-1.0000	-1.0000
	-0.8000	0.6000	-0.8000	2.0000	-0.8000	0.6000	-0.8000	0.0000
Design								
	0.4279	-1.0391	0.2809	0.7126	0.4279	-1.0391	-1.1080	1.3370
X_1'	-0.0193	-0.3197	-1.3211	-0.3556	-0.0193	-0.3197	0.4993	1.8556
	-0.1811	0.8999	0.2484	-1.1704	-0.1811	0.8999	1.1798	-1.6954
Left–singular vectors P'								
	-0.3308	0.2946	-0.3736	0.6594	-0.3308	0.2946	-0.1792	-0.0342
P'	0.0960	-0.1138	0.6024	0.3785	0.0960	-0.1138	-0.5082	-0.4372
	0.0996	-0.3618	-0.1302	0.2927	0.0996	-0.3618	-0.3435	0.7055
Right–singular vectors								
	-0.7099	-0.1307	-0.6921					
Q	0.3508	-0.9177	-0.1865					
	0.6108	0.3752	-0.6973					
Detection probabilities: design X								
	0.4947	0.1828	0.0577					

In particular, discovering alternatives $(\beta - \beta_o)'$ in the direction $\theta_3 = -[0.6921, 0.1865, 0.6973]$ is seen to be unlikely, with power 0.0577. On the other hand, suppose instead that it is critical in context to discover alternatives in the negative orthant of $\mathbb{R}^3$. Then the design X_1 serves to reverse these so that alternatives in the directions of $\{\theta_3, \theta_2, \theta_1\}$ are now detected with probabilities $[0.4947, 0.1828, 0.0577]$, respectively.

Further properties of the design X and its reversal X_1 deserve mention. Observe that $X_0' X_0 = \mathrm{Diag}(n, X'X)$, whereas

$X'X = Q \, \mathrm{Diag}(\kappa_1, \kappa_2, \kappa_3) Q'$ and $X_1' X_1 = Q_R \, \mathrm{Diag}(\kappa_1, \kappa_2, \kappa_3) Q_R'$ where now $Q_R = [q_3, q_2, q_1]$ following the convention that $\{\kappa_1 \geq \kappa_2 \geq \kappa_3\}$ remain ordered. Clearly $V(\hat{\beta} \mid X) = \sigma^2 (X'X)^{-1}$ and $V(\hat{\beta} \mid X_1) = \sigma^2 (X_1' X_1)^{-1}$ differ, where their diagonal elements are listed as variances in Table 5. However, their eigenvalues are identical by construction, as are their $\{A, D, E\}$ efficiency indices as the trace, determinant, and largest eigenvalues of $(X_0' X_0)^{-1}$ under both designs X and X_1.

Table 5 Variances of OLS solutions; eigenvalues of the dispersion matrix $\Gamma \in \{\Sigma_0, \Xi_0\}$ as $\left(X_0' X_0\right)^{-1}$ for the designs $\{X, X_1\}$; and $A, D,$ *and* E efficiencies for these designs

Design Characteristics				
Design	X		X	
Estimates	Variances		Eigenvalues of Γ	
$\hat{\beta}_0$	0.12500	0.12500	3.53636	3.53636
$\hat{\beta}_1$	1.38170	1.83546	0.22694	0.22694
$\hat{\beta}_2$	0.69002	0.26120	0.12500	0.12500
$\hat{\beta}_3$	1.76012	1.73518	0.06854	0.06854
Diagnostic	A	D		E
	3.95684	0.00688		3.53636

Summary and discussion

This study reexamines the concept of directional invariance, or isotropy, for distributions on $\mathbb{R}^k$ having location–scale parameters (δ, Ξ). Powers of tests for $H_0 : \delta = \delta_0$ against $H_0 : \delta \neq \delta_0$ often depend on noncentrality parameters of type $\lambda = (\delta - \delta_0)' \Xi^{-1} (\delta - \delta_0)$. The spectral decomposition of Ξ supports the identification of *directed alternatives* in directions determined by the eigenvectors of Ξ, to encompass the alternatives most likely and least likely in a given study. Powers of these types are independent of direction if and only if $\Xi = \sigma^2 I_k$. Applications are drawn in the use of Hotelling[2] T^2 in multivariate samples, and of *F*-tests in linear models. Case studies are given where this approach leads to the discovery of further insight regarding the natural parameters of a problem.

One concept of directional invariance figures prominently in the *SPC* literature. The following is excerpted from Linna et al.,[10]

"It is well known that the ARL performance of multivariate SPC procedures depends heavily on the covariance structure of the observed data. See, for example, Mason et al.,[11] Further, it has been noted by Pignatiello et al.,[12] that many multivariate procedures, including the χ^2 chart, Hotellings T^2 chart, and most of the multivariate CUSUM charts, are directionally invariant. The performance of the multivariate EWMA chart proposed in Lowry et al.,[13] is also directionally invariant. Lowry et al.,[14] and others also note the directional invariance of many of these multivariate control charting methods. Directional invariance means that the performance of a procedure does not depend on the specific direction in p-space of a shift in the mean vector of the process variables being monitored. Instead, performance of a directionally invariant procedure depends only on the statistical (or Mahalanobis) distance between the in-control mean vector μ_0 and the out-of-control mean vector under consideration, μ_1." (Italics supplied.) That is, $D^2_\Xi(\mu_0, \mu_1) = (\mu_0 - \mu_1)' \Xi^{-1} (\mu_0 - \mu_1)$.

Unfortunately, this notion of directional invariance is grossly misleading, is antithetical to the very concept of invariance as in our Definition 1, at best is a misnomer, and in any event deserves to be clarified in the SPC literature. In fact, such essentials as power functions and ARLs do indeed depend on directions of alternatives, as seen in Section 4.2 as counter examples, unless the model is isotropic.

On the other hand, a disclaimer of Tsui et al.,[15] should be noted: "There is no reason in practice, however, for a shift to $\mu = \mu^*$ to be always considered as important as a shift to $\mu = \mu^{**}$ just because the corresponding values of the noncentrality parameters are equal." Our study represents a substantial elaboration on this point.

A Appendix

A.1 rayleigh quotients.

At issue are variational properties of quadratic forms of type $Q(\upsilon) = \upsilon' \Omega \upsilon / \upsilon' \upsilon$, known as Rayleigh Quotients; see Bellman.[16] Write $\Omega = \Sigma_{i=1}^{k} \kappa_i q_i q_i' = Q D_\kappa Q'$ in its spectral form with $Q \in \mathcal{O}_k$ so that $\{q_i' \Omega q_i = \kappa_i; 1 \le i \le k\}$. Further partition $Q = [Q_1, Q_2, Q_3]$ of orders $\{(k \times (r-1)), (k \times s)\,(k \times d)\}$ with $d = k - r - s + 1$, and partition $D_\kappa = \mathrm{Diag}(D_{\kappa_1}, D_{\kappa_2}, D_{\kappa_3})$ conformably. Then

$$\Omega = Q_1 D_{\kappa_1} Q_1' + Q_2 D_{\kappa_2} Q_2' + Q_3 D_{\kappa_3} Q_3' \qquad (11)$$

where, in particular, $Q_2 = [q_r, ..., q_{r+s}]$ and $D_{\kappa_2} = \mathrm{Diag}(\kappa_r, ..., \kappa_{r+s})$. Essentials follow.

Lemma A.1 Consider the positive definite form $\upsilon' \Omega \upsilon / \upsilon' \upsilon$ with Ω as in expression (11).

(i) Variational properties of $Q(\upsilon)$ as u varies over $\mathbb{R}^k$ are

$$\kappa_k \le \frac{\upsilon' \Omega \upsilon}{\upsilon' \upsilon} \le \kappa_1 \qquad (12)$$

where the lower and upper limits are attained at $\upsilon = q_k$ and $\upsilon = q_1$, respectively.

(ii) Variational properties of $Q(\upsilon)$ as u varies over $S_p(Q_2)$ are

$$\kappa_{r+s} \le \frac{\upsilon' \Omega \upsilon}{\upsilon' \upsilon} \le \kappa_r \qquad (13)$$

where the lower and upper limits are attained at $\upsilon = q_{r+s}$ and $\upsilon = q_r$, respectively.

Proof: Conclusion (i) is given in Bellman,[16] where the limits are attained as given since $\kappa_k = q_k' \Omega q_k$ and $\kappa_1 = q_1' \Omega q_1$ To see conclusion (ii), $\upsilon \in S_p(Q_2)$ may be represented as $\upsilon = Q_2 a$ with $a' = [a_r, ..., a_{r+s}]$. Then for $\upsilon \in S_p(Q_2)$, (11) gives $\upsilon' \Omega \upsilon = a' Q_2' Q_2 D_{\kappa_2} Q_2' Q_2 a = \Sigma_{i=r}^{r+s} \kappa_i a_i^2$. The lower and upper limits for $\upsilon \in S_p(Q_2)$ follow as in conclusion (i) for $\upsilon \in \mathbb{R}^k$, except that now these are attained as given since $q_{r+s}' \Omega q_{r+s} = \kappa_{r+s}$ and $q_r' \Omega q_r = \kappa_r$.

Acknowledgement

References

1. Mahalanobis PC. On the generalised distance in statistics. *Proceedings National Institute of Science India*. 1936;2(1):49–55.

2. Hotelling H. The generalization of Student's ratio. *Ann Math Statist*. 1931;2:360–378.

3. Mauchly JW. Significance test for sphericity of a normal n-variate distribution. *Ann Math Statist*. 1940;11(2):204–209.

4. Rencher AC. Methods of Multivariate Analysis, 2nd ed. Wiley, USA: New York; 2002.

5. Kramer CY, Jensen DR. Fundamentals of Multivariate Analysis, Part I. Inference about Means. *J Quality Technology*. 1969;1:120–133.

6. Hotelling H. Multivariate quality control. In: Eisenhart C, editor. Techniques of Statistical Analysis, McGraw-Hill, USA: New York; 1947. p. 111–184.

7. Quesenberry CP. The multivariate short-run snapshot Q chart. *Quality Engineering*. 2001;13(4):679–683.

8. Williams JD, Woodall WH, Birch JB, et al. Distribution of Hotelling's T² statistic based on sucessive differences covariance matrix estimator. *J Quality Technology*. 2005;38:217–229.

9. Box GEP, Wilson KB. On the experimental attainment of optimum conditions. *J Royal Statist Soc Ser B*. 1951;13(1):1–45.

10. Linna KW, Woodall WH, Busby KL. Performance of multivariate control charts in the presence of measurement error. *J Quality Technology*. 2001;33:349–355.

11. Mason RL, Champ CW, Tracy ND, et al. Assessment of multivariate process control techniques. *J Quality Technology*. 1997;29:140–143.

12. Pignatiello JJ, Runger GC. Comparisons of multivariate CUSUM charts. *J Quality Technology*. 1990;22(3):173–186.

13. Lowry CA, Woodall WH, Champ CW, et al. A multivariate exponentially weighted moving average control chart. *Technometrics*. 1992;34(1):46–53.

14. Lowry CA, Montgomery DC. A review of multivariate control charts. *IIE Transactions*. 1995;27(6):800–810.

15. Tsui KL, Woodall WH. Multivariate control charts based on loss functions. *Sequential Analysis*. 1993;12(1):79–92.

16. Bellman R. Introduction to Matrix Analysis. McGraw-Hill, USA: New York; 1960.