Changepoint Analysis as a Method for Isotonic Inference

CHIHIRO HIROTSU

Meisei University, Japan

KOHEI MARUMO

Nippon-Ginko, Japan

ABSTRACT. Concavity and sigmoidicity hypotheses are developed as a natural extension of the simple ordered hypothesis in normal means. Those hypotheses give reasonable shape constraints for obtaining a smooth response curve in the non-parametric input–output analysis. The slope change and inflection point models are introduced correspondingly as the corners of the polyhedral cones defined by those isotonic hypotheses. Then a maximal contrast type test is derived systematically as the likelihood ratio test for each of those changepoint hypotheses. The test is also justified for the original isotonic hypothesis by a complete class lemma. The component variables of the resulting test statistic have second or third order Markov property which, together with an appropriate non-linear transformation, leads to an exact and very efficient algorithm for the probability calculation. Some considerations on the power of the test are given showing this to be a very promising way of approaching to the isotonic inference.

Key words: concavity, extended max t test, inflection, maximin linear test, non-parametric dose–response analysis, Markov property, sigmoidicity, slope change

1. Introduction

The concavity or convexity hypothesis has naturally been introduced for analysing the age–period–cohort effects model where all those three effects have natural ordering and yet only the second order differences are estimable for each of them (Hirotsu, 1988). The concavity is also a shape constraint typically met in economic models such as utility and production functions. It is often found plausible, for example, to assume that holding land fixed the output of corn rises with the input of seed but with diminishing returns, see Matzkin (1994), for example. A dose–response curve is essentially sigmoidal if the dose range is taken sufficiently large so that we can assume the convexity under the inflection point if the sigmoidicity hypothesis is supported and the inflection point is detected. The convexity assumption is useful not only for obtaining the smooth response curve but also for the low dose extrapolation of the risks in the non-parametric toxicity analysis. Hirotsu & Srivastava (2000) show, for example, how to improve the simultaneous upper bound of the risks under the convexity assumption. In this respect also, the maximal contrast type test statistic is useful since it can point out the inflection point.

Now, suppose we have data $y_1, \ldots, y_K$ from the K independent normal populations $N(\mu_k, \sigma^2),\ k = 1, \ldots, K$. Then a relationship has been demonstrated in Hirotsu (1997) between the simple ordered hypothesis

$$H_1\colon \mu_1 \le \cdots \le \mu_K$$

with at least one inequality strict and the changepoint hypothesis

$$K_1\colon \mu_1 = \cdots = \mu_\tau < \mu_{\tau+1} = \cdots = \mu_K, \quad \text{for some } \tau = 1, \ldots, K-1,$$

where τ is an unknown changepoint and the null hypothesis is

$$H_0\colon \mu_1 = \cdots = \mu_K$$

for both cases. It is simply that a set of component hypotheses of K_1 indexed by τ forms the $K-1$ corners of the polyhedral cone defined by H_1. Assuming σ^2 to be known tentatively the likelihood ratio test for K_1 is easily obtained as

$$\max_{\tau=1,\ldots,K-1} \left(\frac{1}{\tau} + \frac{1}{K-\tau}\right)^{-1/2} \frac{1}{\sigma}\left(\frac{Y_\tau^*}{K-\tau} - \frac{Y_\tau}{\tau}\right) > c, \tag{1}$$

where $Y_\tau = y_1 + \cdots + y_\tau$, $Y_\tau^* = y_{\tau+1} + \cdots + y_K$ and c is chosen to meet the required significance level; see Sen & Srivastava (1975) for the derivation of the statistic. The test statistic is also interpreted as the standardized maximum of the projections of the efficient score vector onto the $K-1$ corners of the polyhedral cone, where an efficient score vector is defined as the derivative of the log likelihood with respect to the parameter μ_k, $k = 1, \ldots, K$, and evaluated at the null hypothesis. On the other hand a complete class of tests for the ordered alternative H_1 is given by all the tests that are increasing in every element of those $(K-1)$ projections and with convex acceptance regions, see Hirotsu (1982) for details. Then it happens that the likelihood ratio test statistic for K_1 has been independently proposed and justified also in the stream of the isotonic inference as a useful test statistic. In this sense it should be compared with the maximal contrast type test by Williams (1971) and its modification by Marcus (1976). Actually the test statistic (1) has been called the max t and shown, as compared with other methods, to have high power in the wide range of the ordered alternatives specified by H_1; see Hirotsu (1979) and Hirotsu *et al.* (1992) for details. The max t statistic in the unbalanced one-way layout setting is also introduced in Hirotsu *et al.* (1992). In this paper we extend the relationship to more general isotonic hypotheses which will be useful, for example, as shape constraints for a non-parametric dose–response analysis. It is interesting and useful to find out the corner models to figure out those generalized isotonic hypotheses.

In section 2 of the present paper we first introduce the concavity hypothesis as one of those reasonable shape constraints and derive a slope change model as its corner model. Then the likelihood ratio test statistic for the slope change model is developed as an extension of max t and shown to be an appropriate test statistic also for the concavity hypothesis. Similarly the sigmoidicity hypothesis and a corresponding inflection point model are demonstrated in section 3 as well as the extended max t test derived as a likelihood ratio test for the latter model. Section 4 is devoted to the distribution theory of those extended max t tests and a very efficient recurrence formula for the level probability and also for the power calculation is given. It is based on the Markov property of the component variables of the test statistic and considered as a natural extension of the approach taken by Hawkins (1977) and Worsley (1986) for the simple changepoint hypothesis. In this case, however, the Markov properties are second and third order differently from the first order in the previous papers and a non-linear transformation introduced in section 5 is inevitable for reducing the number of the cut points for the conditioning variables in the recurrence formula. Some considerations on the power are given in section 6. It is first argued that the maximal angles of the polyhedral cones defined by the concavity and sigmoidicity hypotheses are smaller than those defined by the simple order suggesting the appropriateness of the extended max t as compared with the classical max t. Then the maximin linear test is introduced for comparisons of powers for each of the concavity and sigmoidicity hypotheses. The comparisons are in favour of those extended max t tests. All the theories are given for the normal model but they can be applied asymptotically also to the rank or the binomial data and have much wider applications. An application to the binomial data y_i is given in section 7. It should be noted that since the extended max t tests are, like max t, based on the weighted sum of the y_i the Lindeberg condition for asymptotic normality is more easily met than that for each y_i, see app. 3 of Hirotsu & Srivastava (2000).

2. The concavity hypothesis and a slope change model

2.1. Mathematical formulation

Suppose we have data from a one-way layout

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, \ldots, K; \quad j = 1, \ldots, n_i,$$

where the ε_{ij} are independently distributed as $N(0, \sigma^2)$ and we assume tentatively σ^2 to be known. Suppose we have an explanatory variable x_i at the level i, $x_1 \le \cdots \le x_K$, then a concavity hypothesis is defined as

$$H_2\colon \frac{\mu_2 - \mu_1}{x_2 - x_1} \ge \frac{\mu_3 - \mu_2}{x_3 - x_2} \ge \cdots \ge \frac{\mu_K - \mu_{K-1}}{x_K - x_{K-1}}$$

with at least one inequality strict. This is an extension of the monotone hypothesis in the first order differences of means introduced in Hirotsu (1986) and goes back to the previous situation if we take x_is equally spaced. Now we can give lemma 1 asserting a relationship between the concavity hypothesis and a slope change model.

Lemma 1
Each component of a slope change model defined by

$$K_2(\tau)\colon \begin{cases} \dfrac{\mu_2 - \mu_1}{x_2 - x_1} = \cdots = \dfrac{\mu_{\tau+1} - \mu_\tau}{x_{\tau+1} - x_\tau} = \beta, \\[2ex] \dfrac{\mu_{\tau+2} - \mu_{\tau+1}}{x_{\tau+2} - x_{\tau+1}} = \cdots = \dfrac{\mu_K - \mu_{K-1}}{x_K - x_{K-1}} = \beta^* \ (< \beta) \end{cases}$$

and indexed by $\tau = 1, \ldots, K-2$ forms a set of $K-2$ corner vectors of the polyhedral cone defined by H_2.

Proof. The concavity hypothesis H_2 can be written in the matrix notation as

$$L_K^{*\prime} \boldsymbol{\mu} \ge \mathbf{0} \tag{2}$$

with

$$L_K^{*\prime} = \begin{bmatrix} \frac{1}{x_1 - x_2} & \frac{1}{x_2 - x_1} + \frac{1}{x_3 - x_2} & & \frac{1}{x_2 - x_3} & 0 & \cdots & 0 \\ & & \cdots & & & & \\ 0 & \cdots & 0 & \frac{1}{x_{K-2} - x_{K-1}} & \frac{1}{x_{K-1} - x_{K-2}} + \frac{1}{x_K - x_{K-1}} & & \frac{1}{x_{K-1} - x_K} \end{bmatrix}_{K-2 \times K}.$$

Since there are two additional degrees of freedom to define $\boldsymbol{\mu}$ satisfying (2) the restriction

$$B' \boldsymbol{\mu} = \mathbf{0} \tag{3}$$

is imposed, where

$$B = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_K \end{pmatrix}'.$$

It is easy to verify $B'L_K^* = 0$ and we can add restriction (3) without violating (2). Then all the $\boldsymbol{\mu}$ satisfying (2) and (3) can be expressed as such $\boldsymbol{\mu}$ satisfying

$$\begin{pmatrix} B' \\ L_K^{*\prime} \end{pmatrix} \boldsymbol{\mu} = \begin{pmatrix} \mathbf{0} \\ \boldsymbol{h} \end{pmatrix}$$

with some $\boldsymbol{h} \geq \mathbf{0}$. Then for all those $\boldsymbol{\mu}$ we have an expression

$$\begin{aligned} \boldsymbol{\mu} &= \{B(B'B)^{-1}B' + L_K^*(L_K^{*\prime}L_K^*)^{-1}L_K^{*\prime}\}\boldsymbol{\mu} \\ &= L_K^*(L_K^{*\prime}L_K^*)^{-1}\boldsymbol{h}. \end{aligned}$$

The first equality holds since $\Pi_B = B(B'B)^{-1}B'$ and $\Pi_{L_K^*} = L_K^*(L_K^{*\prime}L_K^*)^{-1}L_K^{*\prime}$ are the projection matrices of rank 2 and $K-2$ and orthogonal each other. Thus any $\boldsymbol{\mu}$ satisfying (2) and (3) can be expressed by the positive linear combination of the columns of $L_K^*(L_K^{*\prime}L_K^*)^{-1}$ like $\boldsymbol{\mu} = L_K^*(L_K^{*\prime}L_K^*)^{-1}(L_K^{*\prime}\boldsymbol{\mu})$. Conversely it is obvious that every $\boldsymbol{\mu}$ expressed by $L_K^*(L_K^{*\prime}L_K^*)^{-1}\boldsymbol{h}, \boldsymbol{h} \geq \mathbf{0}$ satisfies the restrictions (2) and (3). It implies that $K-2$ columns of $L_K^*(L_K^{*\prime}L_K^*)^{-1}$ give the corner vectors of the polyhedral cone defined by (2) and (3). Excluding the restriction (3) we have an expression for $\boldsymbol{\mu}$ satisfying H_2 like

$$\boldsymbol{\mu} = B(\eta_0, \eta_1)' + L_K^*(L_K^{*\prime}L_K^*)^{-1}\boldsymbol{h}, \quad \boldsymbol{h} \geq \mathbf{0} \tag{4}$$

with η_0 and η_1 arbitrary regression coefficients. Now we give an explicit form of $L_K^*(L_K^{*\prime}L_K^*)^{-1}$ to complete the proof. First we rewrite the model $K_2(\tau)$ in the linear form like

$$K_2(\tau)\colon \boldsymbol{\mu} = (B \ \ \boldsymbol{b}_\tau)\begin{pmatrix} \eta_0 \\ \eta_1 \\ \eta \end{pmatrix} = B(\eta_0, \eta_1)' + \boldsymbol{b}_\tau \eta_\tau \tag{5}$$

with $\eta_\tau = \beta - \beta^* > 0$ and

$$\boldsymbol{b}_\tau = (\Pi_B - I)(0, \ldots 0, x_{\tau+2} - x_{\tau+1}, \ldots, x_K - x_{\tau+1})'.$$

Then it is easy to verify $L_K^{*\prime}\boldsymbol{b}_\tau$ to be equal to $(0 \cdots 010 \cdots 0)'$ with a unit element as its τth component and this, together with the relation $B'\boldsymbol{b}_\tau = \mathbf{0}$, implies the equality

$$L_K^*(L_K^{*\prime}L_K^*)^{-1} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{K-2}). \tag{6}$$

Then by comparing (4) and (5) we see that every model $K_2(\tau)$ indexed by $\tau = 1, \ldots, K-2$ forms $K-2$ corners of H_2.

A brief sketch of the cone and its standardized corner vectors are given in Fig. 1 for $K = 5$ and equal spacing case. The direction of second order polynomial $\boldsymbol{\mu}_q = (-2 \ 1 \ 2 \ 1 \ -2)'$ is shown to be located inside the cone.

2.2. *Test statistic*

In the expression (5) the generalized least squares estimator of η_τ is obtained as

$$S_\tau^* = M_\tau^{-1}\boldsymbol{b}_\tau'\Omega^{-1}\{I - B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}\}\bar{\boldsymbol{y}}$$

with variance M_τ^{-1}, where $\bar{\boldsymbol{y}} = (\bar{y}_{1.}, \ldots, \bar{y}_{K.})'$ is the vector of means with dispersion matrix $\Omega = \mathrm{diag}(\sigma^2/n_i)$ a diagonal matrix with σ^2/n_i as its ith diagonal element and $M_\tau = \boldsymbol{b}_\tau'\Omega^{-1}\boldsymbol{b}_\tau - \boldsymbol{b}_\tau'\Omega^{-1}B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}\boldsymbol{b}_\tau$. Then the likelihood ratio test statistic for the null hypothesis

$$H_B\colon \boldsymbol{\mu} = B(\eta_0, \eta_1)'$$

against the one-sided slope change alternative $K_2(\tau)$ with $\eta > 0$ at given τ is obtained as $S_\tau = M_\tau^{1/2}S_\tau^*$. Actually S_τ is the square root of minus twice the log likelihood ratio but it will be standardized corner vectors

$$\begin{cases} \mu_1^\dagger = (\ -2 \ \ 2 \ \ 1 \ \ 0 \ -1 \ ) \ /\sqrt{10} \\ \mu_2^\dagger = (\ -4 \ \ 1 \ \ 6 \ \ 1 \ -4 \ ) \ /\sqrt{70} \\ \mu_3^\dagger = (\ -1 \ \ 0 \ \ 1 \ \ 2 \ -2 \ ) \ /\sqrt{10} \end{cases}$$

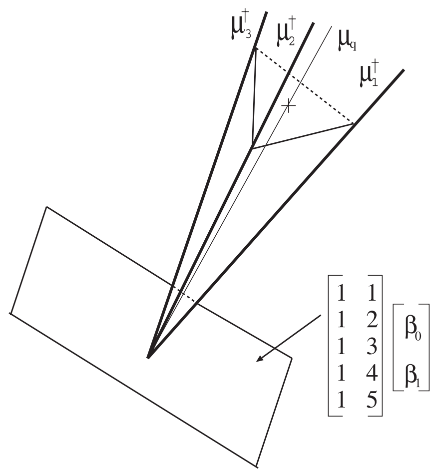

Fig. 1. Polyhedral cone defined by (2) and its corner vectors with $K = 5$ and equal spacing.

natural to call it the likelihood ratio statistic concerning the one-sided alternative. The likelihood ratio test statistic for the slope change model is then obtained by taking the maximum of S_τ over $\tau = 1, \ldots, K-2$,

$$S = \max S_\tau,$$

thus giving a natural extension of max t of (1). By virtue of the relation (6) we have a very convenient expression of $\boldsymbol{s} = (S_1, \ldots, S_{K-2})'$ like

$$\boldsymbol{s} = \mathrm{diag}(M_\tau^{-1/2})(L_K^{*\prime} L_K^*)^{-1} L_K^{*\prime} \Omega^{-1} \{\bar{\boldsymbol{y}} - \hat{E}_0(\bar{\boldsymbol{y}})\}, \tag{7}$$

where $\hat{E}_0(\bar{\boldsymbol{y}}) = B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}\bar{\boldsymbol{y}}$ is the maximum likelihood estimator of the mean vector under the null hypothesis H_B. Thus each component of the statistic $\boldsymbol{s}$ is understood as the standardized projection of the efficient score vector evaluated at the null hypothesis H_B onto the corner vector of the polyhedral cone defined by H_2 and then the likelihood ratio test is supported by a compete class lemma given in Hirotsu (1982) to be an appropriate test also for the concavity hypothesis H_2. It is shown by power comparisons in section 6 that the test has some advantage over the Abelson & Tukey (1963) type maximin linear test against the concavity hypothesis.

It should be noted that the convexity hypothesis can be dealt with just by inverting the sign of the test statistic.

3. The sigmoidicity hypothesis and an inflection point model

We go one step ahead of H_2 by considering the ordered hypothesis in the second order differences,

$$H_3\colon \frac{1}{x_3 - x_1}\left(\frac{\mu_3 - \mu_2}{x_3 - x_2} - \frac{\mu_2 - \mu_1}{x_2 - x_1}\right) \geq \cdots \geq \frac{1}{x_K - x_{K-2}}\left(\frac{\mu_K - \mu_{K-1}}{x_K - x_{K-1}} - \frac{\mu_{K-1} - \mu_{K-2}}{x_{K-1} - x_{K-2}}\right). \tag{8}$$

Then we can show each component of the model defined by

$$K_3(\tau)\colon \boldsymbol{\mu} = [C\boldsymbol{c}_\tau](\eta_0, \eta_1, \eta_2, \eta)', \quad \eta > 0,$$

and indexed by $\tau = 1, \ldots, K-3$ with

$$C = \begin{pmatrix} 1 & x_1 & x_1^2 \\ \vdots & \vdots & \vdots \\ 1 & x_K & x_K^2 \end{pmatrix},$$

and

$$\boldsymbol{c}_\tau = (\Pi_c - I)(0 \cdots 0 \ (x_{\tau+3} - x_{\tau+1})(x_{\tau+3} - x_{\tau+2}) \cdots (x_K - x_{\tau+1})(x_K - x_{\tau+2}))',$$
$$\Pi_c = C(C'C)^{-1}C'$$

forms a set of $K-3$ corner vectors of the polyhedral cone defined by H_3, see appendix for a proof. It should be noted that the model $K_3(\tau)$ is composed of two segments of the second order polynomials having two common values at $x = x_{\tau+1}$ and $x_{\tau+2}$, see Fig. 2. This is in contrast to the slope change model $K_2(\tau)$ where two segments of the linear equations have only one common value at $x = x_{\tau+1}$. The model $K_3(\tau)$ suggests a change of response curve from convex to concave between the two points $x_{\tau+1}$ and $x_{\tau+2}$ and may be called an inflection point model and then we call H_3 a sigmoidicity hypothesis. It is not exactly the same with Schmoyer's (1984) definition of sigmoidicity which is the unimodal hypothesis of the slopes of the subsequent segments but has a close relationship.

By similar arguments to section 2.2 the likelihood ratio test statistic for the inflection point model is obtained as the maximal component of

$$\boldsymbol{t} = \mathrm{diag}(N_\tau^{-1/2})(Q_K^{*\prime}Q_K^*)^{-1}Q_K^{*\prime}\Omega^{-1}\{\bar{\boldsymbol{y}} - \hat{E}_0(\bar{\boldsymbol{y}})\},$$

where N_τ is M_τ with $\boldsymbol{b}_\tau$ and B replaced by $\boldsymbol{c}_\tau$ and C, respectively, $Q_K^{*\prime}$ a $K-3 \times K$ coefficient matrix in expressing the inequality (8) like $Q_K^{*\prime}\boldsymbol{\mu} \geq \mathbf{0}$ and $\hat{E}_0(\bar{\boldsymbol{y}}) = C(C'\Omega^{-1}C)^{-1}C'\Omega^{-1}\bar{\boldsymbol{y}}$ in this case. The explicit form of $Q_K^{*\prime}$ is given in the appendix. The statistic is again appropriate for the sigmoidicity hypothesis by virtue of the complete class lemma.

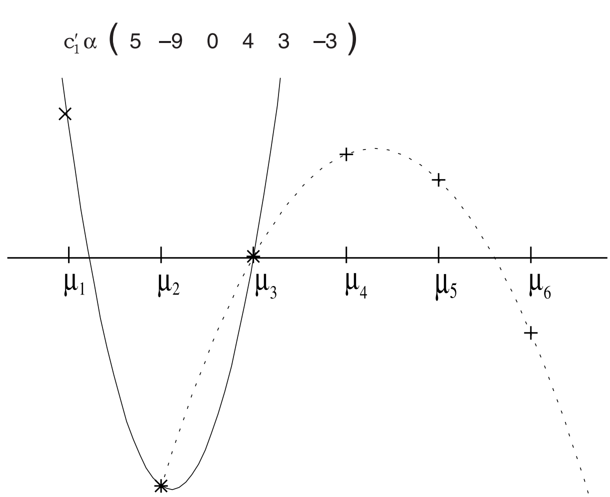

Fig. 2. Sketch of $\eta\boldsymbol{C}_\tau'$ with $K = 6$, $\tau = 1$ and equal spacing.

4. Distribution theory of the maximal components of *s* and *t*

The distribution theory for $\boldsymbol{s}$ and $\boldsymbol{t}$ goes almost parallel and we mainly deal with $\boldsymbol{s}$ here and only the result is stated briefly for $\boldsymbol{t}$.

By a simple matrix algebra and as shown generally in Hirotsu (1982) the statistic $\boldsymbol{s}$ of (7) can be written as

$$\boldsymbol{s} = \mathrm{diag}(M_\tau^{-1/2})(L_K^{*\prime}\Omega L_K^*)^{-1}L_K^{*\prime}\bar{\boldsymbol{y}}$$

and then the covariance matrix is obtained simply as

$$\mathrm{var}(\boldsymbol{s}) = \mathrm{diag}(M_\tau^{-1/2})(L_K^{*\prime}\Omega L_K^*)^{-1}\ \mathrm{diag}(M_\tau^{-1/2}). \tag{9}$$

By virtue of the form of $L_K^{*\prime}$ and diagonal matrix Ω this is an inverse matrix of a penta-diagonal matrix and it is easy to show that for any partition $(\boldsymbol{s}_1', \boldsymbol{s}_2')'$ of $\boldsymbol{s}$ the conditional distribution of $\boldsymbol{s}_1$ given $\boldsymbol{s}_2$ depends only on the first two elements of $\boldsymbol{s}_2$ implying the second order Markov property in the sequence $S_1, \ldots, S_{K-2}$ of the components of $\boldsymbol{s}$. By this Markov property we have a simple recurrence formula for the joint probability

$$P_i(s_o) = \mathrm{pr}\{S_1 < s_o, \ldots, S_{K-2} < s_o\},$$

where the index i takes 0 or 1 according to the null or the alternative distribution. Then the p value for the observed maximum s_o is obtained as

$$p = \mathrm{pr}\{\max S_\tau \geq s_o | H_0\} = 1 - P_0(s_o).$$

For the recurrence formula define the conditional probability

$$F_{\tau\tau+1}(s_o|S_\tau, S_{\tau+1}) = \mathrm{pr}\{S_1 < s_o, \ldots, S_{\tau+1} < s_o | S_\tau, S_{\tau+1}\}, \quad \tau = 1, \ldots, K-1,$$

where for convenience we introduce S_{K-1} and S_K which are defined to be zero so that $P_i(s_o) = F_{K-1K}(s_o|S_{K-1}, S_K)$ is an unconditional probability.

Starting from the initial function

$$F_{12}(s_o|S_1, S_2) = \begin{cases} 1, & S_1 < s_o, S_2 < s_o\ , \\ 0, & \text{otherwise} \end{cases}$$

we can calculate $F_{\tau\tau+1}$ recursively by a single numerical integration with respect to S_τ. We state the formula in lemma 2.

Lemma 2 *Recurrence formula for* $F_{\tau\tau+1}$
For $\tau = 1, \ldots, K-2$ *we have*

$$F_{\tau+1\tau+2}(s_o|S_{\tau+1}, S_{\tau+2}) = \begin{cases} \int_{-\infty}^{s_o} & F_{\tau\tau+1}(s_o|S_\tau, S_{\tau+1}) f_{\tau|\tau+1,\tau+2}\, dS_\tau, S_{\tau+2} < s_o\ , \\ 0, & \text{otherwise,} \end{cases} \tag{10}$$

where $f_{\tau|\tau+1,\tau+2} = f_{\tau|\tau+1,\tau+2}(S_\tau|S_{\tau+1}, S_{\tau+2})$ *is the conditional probability density function of* S_τ *given* $S_{\tau+1}$ *and* $S_{\tau+2}$.

Proof. By the law of total probability we have

$$F_{\tau+1\tau+2}(s_o|S_{\tau+1}, S_{\tau+2}) = \int \mathrm{pr}(S_1 < s_o, \ldots, S_{\tau+2} < s_o | S_\tau, S_{\tau+1}, S_{\tau+2}) f_{\tau|\tau+1,\tau+2}\, dS\tau.$$

If the inequality $S_{\tau+2} < s_o$ is satisfied then we can discard the same inequality in the integrand and get

$$F_{\tau+1\tau+2}(s_o|S_{\tau+1}, S_{\tau+2}) = \int F_{\tau\tau+1}(s_o|S_\tau, S_{\tau+1}) f_{\tau|\tau+1,\tau+2}\, dS\tau$$

applying the second order Markov property in $S'_\tau s$. If $S_{\tau+2} \geq s_o$ obviously

$$F_{\tau+1\tau+2}(s_o|S_{\tau+1}, S_{\tau+2}) = 0.$$

For the conditional density $f_{\tau|\tau+1,\tau+2}$ let ρ_{ij} denote the (i, j) element of (9), then the conditional distribution of S_τ given $S_{\tau+1}$ and $S_{\tau+2}$ is a normal with mean

$$\mu_{s\tau} + (\rho_{\tau\tau+1}, \rho_{\tau\tau+2}) \begin{pmatrix} 1 & \rho_{\tau+1\tau+2} \\ \rho_{\tau+1\tau+2} & 1 \end{pmatrix}^{-1} (S_{\tau+1} - \mu_{s\tau+1}, S_{\tau+2} - \mu_{s\tau+2})' \tag{11}$$

and variance

$$1 - (\rho_{\tau\tau+1}, \rho_{\tau\tau+2}) \begin{pmatrix} 1 & \rho_{\tau+1\tau+2} \\ \rho_{\tau+1\tau+2} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{\tau\tau+1} \\ \rho_{\tau\tau+2} \end{pmatrix} \tag{12}$$

for $\tau = 1, \ldots, K-4$, where $\mu_{s\tau}$ is the τth component of $E(\boldsymbol{s})$. The last two steps of the recurrence formula need some caution but we can deal with it simply by extending the definition of (ρ_{ij}) up to $1 \leq i, j \leq K$ by

$$\rho_{ij} = \begin{cases} \rho_{ij}, & 1 \leq i, j \leq K-2\ , \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Then the formulae (11) and (12) for the conditional density $f_{\tau|\tau+1,\tau+2}$ can be extended to $\tau = K-3$ and $K-2$ as it is. It is easy, for example, to see that by the definition (13) we have unconditional normal density $N(\mu_s \eta_{K-2}, 1)$ for $f_{K-2|K-1,K}$. Thus we can obtain P_i only by the use of single integration recursively. Finally the difference between P_0 and P_1 is only that all the $\mu_{s\tau}$ vanish in the calculation of the conditional density for P_0.

Similarly we have the third order Markov property for $\boldsymbol{t}$ which brings forth a recurrence formula for the joint probabilities

$$P_i(t_o) = \text{pr}\{T_1 < t_o, \ldots, T_{K-3} < t_o\}, \quad i = 0, 1$$

based on the conditional probability

$$F_{\tau\tau+1\tau+2}(t_o|T_\tau, T_{\tau+1}, T_{\tau+2}) = \text{pr}\{T_1 < t_o, \ldots, T_{\tau+2} < t_o|T_\tau, T_{\tau+1}, T_{\tau+2}\}.$$

The recurrence formula starts from the initial function

$$F_{123}(t_o|T_1, T_2, T_3) = \begin{cases} 1, & T_1 < t_o, T_2 < t_o, T_3 < t_o, \\ 0, & \text{otherwise} \end{cases}$$

and renewed by the formula

$$F_{\tau+1\tau+2\tau+3}(t_o|T_{\tau+1}, T_{\tau+2}, T_{\tau+3})$$
$$= \begin{cases} \int_{-\infty}^{t_o} & F_{\tau\tau+1\tau+2}(t_o|T_\tau, T_{\tau+1}, T_{\tau+2}) f_{\tau|\tau+1,\tau+2,\tau+3}(T_\tau|T_{\tau+1}, T_{\tau+2}, T_{\tau+3})\, dT_\tau, T_{\tau+3} < t_o, \\ 0, & \text{otherwise.} \end{cases} \tag{14}$$

up to $\tau = K-3$, where $f_{\tau|\tau+1,\tau+2,\tau+3}$ is the conditional normal density and we define $T_{K-2} = T_{K-1} = T_K = 0$. Again only a single integration is necessary although the conditioning is now 3-dimensional.

Finally when σ^2 is unknown but available its unbiased estimate $\hat{\sigma}^2$ which is distributed as a constant times chi-squared variable independently of $\bar{\boldsymbol{y}}$ the formulae (10) and (14) can be easily extended. It is only necessary to use the Studentized statistics by replacing σ^2 by $\hat{\sigma}^2$ in calculating the test statistics $\boldsymbol{s}$ and $\boldsymbol{t}$. Then in calculating F_{K-1K} and $F_{K-2K-1K}$, replace s_o and t_o

by $s_o(\hat{\sigma}/\sigma)$ and $t_o(\hat{\sigma}/\sigma)$, respectively, and take the expectation of the results with respect to the distribution of $\hat{\sigma}/\sigma$, which is a constant times chi variable. For example, denoting the Studentized statistic by $S_\tau^\dagger$ and the chi variable with the degrees of freedom v by χ_v the formula (10) can be rewritten as

$$\int_{-\infty}^{s_o\chi_v/\sqrt{v}} F_{\tau\tau+1}(s_o\chi_v/\sqrt{v}|S_\tau^\dagger, S_{\tau+1}^\dagger) f_{\tau|\tau+1,\tau+2}\, dS_\tau^\dagger$$

and then we obtain $F_{K-1K}(s_o\chi_v/\sqrt{v}|S_{K-1}^\dagger, S_K^\dagger)$ for each χ_v similarly to lemma 2. Finally we obtain the joint probability

$$\begin{aligned} P_i(s_o) &= \mathrm{pr}\{S_1^\dagger < s_o, \ldots, S_{K-2}^\dagger < s_o\} \\ &= \int_0^\infty F_{K-1K}(s_o\chi_v/\sqrt{v}|S_{K-1}^\dagger, S_K^\dagger) g(\chi_v)\, d\chi_v, \end{aligned}$$

where $g(\chi_v)$ is the probability density function of the χ-distribution with the degrees of freedom v. Thus there is no numerical difficulty in dealing with the unknown variance case.

5. An efficient execution of the recurrence formula

In executing the numerical integration of (10) with respect to S_τ it is not possible to have values of the integrand $F_{\tau\tau+1}(u_o|S_\tau, S_{\tau+1})$ beforehand at a small number of points $(S_\tau, S_{\tau+1})$ most convenient for the integration, since an efficient distribution of the points for a numerical integration is usually given only after knowing the shape of the integrand. A naïve method therefore requires to evaluate $F_{\tau\tau+1}$ at a large number of points $(S_\tau, S_{\tau+1})$ and to interpolate for the other points. The method should, however, be very inefficient and even infeasible for the case of the inflection point model where the 3-dimensional conditioning is required. The method also suffers from the errors induced by the interpolation. In the following we propose a very efficient algorithm based on the transformation of the variables S_τ, which will require the calculation of $F_{\tau\tau+1}$ at only $64 \sim 128$ equidistant evaluation points for each transformed variable and in the integration step use only those pre-calculated values of $F_{\tau\tau+1}$ avoiding the interpolation process.

First in the integration of (10) we replace $-\infty$ by a sufficiently large number $-C$ and convert the range of integration into [0,1] by a linear transformation

$$S_\tau^* = (S_\tau + C)/(s_o + C)$$

for each of $\tau = 1, \ldots, K-2$. Then we employ a non-linear transformation

$$S_\tau^* = \varphi(\upsilon_\tau), \quad 0 \le \upsilon_\tau \le 1, \quad \tau = 1, \ldots, K-2$$

with

$$\varphi(\upsilon) = \frac{1}{2} + (2\upsilon - 1)\left\{\frac{1}{2} + \omega + 3\omega^2 + 3\omega^3\right\}, \ \omega = \upsilon(\upsilon - 1)$$

to obtain the recurrence formula in υ_τ

$$F^*_{\tau+1\tau+2}(\upsilon_{\tau+1}, \upsilon_{\tau+2}) = (s_o + C)\int_0^1 F^*_{\tau\tau+1}(\upsilon_\tau, \upsilon_{\tau+1}) f^*_{\tau|\tau+1,\tau+2}(\upsilon_\tau|\upsilon_{\tau+1}, \upsilon_{\tau+2})\varphi'(\upsilon_\tau)d\upsilon_\tau,$$

where φ' is the derivative of φ and

$$\begin{aligned} & f^*_{\tau|\tau+1,\tau+2}(\upsilon_\tau|\upsilon_{\tau+1}, \upsilon_{\tau+2}) \\ &\quad = f_{\tau|\tau+1,\tau+2}\{(s_o + C)\varphi(\upsilon_\tau) - C|(s_o + C)\varphi(\upsilon_{\tau+1}) - C, (s_o + C)\varphi(\upsilon_{\tau+2}) - C\}. \end{aligned}$$

By this transformation a singularity at the border of the integration resolves and we can perform the integration by a simple trapezoidal rule with common evaluation points

$$\upsilon_\tau = I/n, \quad I = 0, 1, \ldots, n \text{ for each } \upsilon_\tau, \quad \tau = 1, \ldots, K-2,$$

see Laurie (1996) for the details of this non-linear transformation.

Now starting from the initial function

$$F^*_{12}(\upsilon_1, \upsilon_2) = \begin{cases} 1, & 0 \le \upsilon_1 < 1,\ 0 \le \upsilon_2 < 1\ , \\ 0, & \text{otherwise,} \end{cases}$$

we proceed recursively by the formula

$$F^*_{\tau+1\tau+2}\left(\frac{I}{n}, \frac{J}{n}\right) = (s_0 + C) \sum_{H=0}^{n} \left\{ F^*_{\tau\tau+1}\left(\frac{H}{n}, \frac{I}{n}\right) f^*_{\tau|\tau+1,\tau+2}\left(\frac{H}{n}\middle|\frac{I}{n}, \frac{J}{n}\right) \varphi'\left(\frac{H}{n}\right) \frac{1}{n} \right\},$$
$$I, J = 0, 1, \ldots, n \tag{15}$$

until $\tau = K-2$, where F^*_{K-2K-1} should be calculated for $I = 0, \ldots, n$ retaining $J = 0$ and F^*_{K-1K} should be calculated only once at $I = J = 0$ with $f^*_{K-2|K-1,K}$ an unconditional density function. It should be noted that we are avoiding the interpolation process. Obviously the function $\varphi(H/n)$ and $\varphi'(H/n)$, $H = 0, 1, \ldots, n$, are common for each step and should be calculated only once. Since for usual purposes the number n of evaluation points can be 64 for the second and 128 for the third order cases the formula (15) provides a very simple and efficient algorithm for evaluating the required joint probability.

It is very easy to write down the formula for the inflection point model based on the third order Markov property and it is omitted here.

6. Power comparisons

The max t test has been verified to keep high powers for the wide range of the restricted alternative in case of the simple ordered hypothesis. On the other hand it has been pointed out that those maximal contrast type tests will not be so useful if the maximal angle of the polyhedral cone is large, see Robertson *et al.* (1988, sect. 4.2–4.4). In particular Abelson & Tukey (1963) type maximin linear test is said to be useful only for the number of levels K under 5 in the simple ordered hypothesis case. We therefore show in Table 1 that the maximal angles of the polyhedral cones treated here are much smaller than those of the simple ordered hypothesis. It is a simple algebra to show that the cosine of the maximum angles are $1/(K-1)$ and $2/(K-1)$ respectively for the monotonicity and convexity hypotheses. This suggests that the maximal contrast type tests introduced here are even more appropriate than the max t test in the simple ordered hypothesis. It suggests, however, the maximin linear test might do also as well. For $K = 6$ and 8 and under equal spacing and equal sample sizes the maximin linear tests are therefore searched for on the corners, edges and faces of each of the polyhedral cones

Table 1. *Comparison of the maximal angles of the three types of polyhedral cones under the equal spacing as expressed by the value of cosine*

	K								
	2	3	4	5	6	7	8	9	10
Monotonicity	1	0.50	0.33	0.25	0.20	0.17	0.14	0.13	0.11
Convexity	—	1	0.67	0.50	0.40	0.33	0.29	0.25	0.22
Sigmoidicity	—	—	1	0.75	0.60	0.50	0.43	0.38	0.33

Table 2. *The coefficients of the two types of linear tests in case of the equal spacing and equal sample size*

	k: 1	2	3	4	5	6	7	8
(a) Maximin linear test								
Convexity	−0.5773	0.2829	0.2944	0.2944	0.2829	−0.5773		
	−0.6108	0.1673	0.2036	0.2399	0.2399	0.2036	0.1673	−0.6108
Sigmoidicity	2	−3	−1	1	3	−2		
	5	−5	−3	−1	1	3	5	−5
(b) Polynomial test								
Convexity	−5	1	4	4	1	−5		
(Quadratic)	−7	−1	3	5	5	3	−1	−7
Sigmoidicity	5	−7	−4	4	7	−5		
(Cubic)	7	−5	−7	−3	3	7	5	−7

Table 3. *Power comparisons of the maximal contrast tests and maximin linear test (equal spacing, equal sample size and σ^2 known)*

Corner and quadratic or cubic configuration								*Maximal contrast*	*Maximin linear*	*Polynomial coefficients*
(1) Convexity hypothesis										
−10	8	5	2	−1	−4			0.698	0.657	0.615
−20	2	24	11	−2	−15			0.721	0.657	0.747
−15	−2	11	24	2	−20			0.721	0.657	0.747
−4	−1	2	5	8	−10			0.698	0.657	0.615
quadratic								0.747	0.751	0.790
−7	4	3	2	1	0	−1	−2	0.674	0.623	0.535
−70	−5	60	41	22	3	−16	−35	0.702	0.626	0.696
−35	−10	15	40	23	6	−11	−28	0.710	0.623	0.755
−28	−11	6	23	40	15	−10	−35	0.710	0.623	0.755
−35	−16	3	22	41	60	−5	−70	0.702	0.626	0.696
−2	−1	0	1	2	3	4	−7	0.674	0.623	0.535
quadratic								0.737	0.726	0.790
(2) Sigmoidicity hypothesis										
3	−3	−4	0	9	−5			0.730	0.707	0.697
15	−19	−18	18	19	−15			0.748	0.737	0.774
5	−9	0	4	3	−3			0.730	0.707	0.697
cubic								0.763	0.780	0.790
7	−9	−3	1	3	3	1	−3	0.705	0.665	0.621
21	−17	−24	0	13	15	6	−14	0.730	0.689	0.745
7	−4	−8	−5	5	8	4	−7	0.735	0.690	0.776
14	−6	−15	−13	0	24	17	−21	0.730	0.689	0.745
3	−1	−3	−3	−1	3	9	−7	0.705	0.665	0.621
cubic								0.751	0.748	0.790

according to Abelson & Tukey (1963) and results are shown in Table 2. Then we compare the powers of the extended max t tests and the maximin linear tests in Table 3 assuming equal sample sizes and σ^2 known. We add in the comparisons the linear tests with coefficients for the quadratic and cubic patterns, which seem to be useful for the concavity and sigmoidicity hypotheses, respectively. The upper percentiles of the extended max t tests have been obtained by solving the equation for the p-value calculation conversely, where the computation is somewhat hard for $K = 8$ of the sigmoidicity hypothesis and the recurrence formula based on the non-linear transformation of variables is essential.

The powers are compared in the directions of the corner vectors and also the quadratic or cubic configuration, where the noncentrality parameter $n \sum(\mu_i - \bar{\mu})^2/\sigma^2$ is fixed at 6 so that powers are around 0.70. It should be noted that the polynomial type test is the most powerful test against the corresponding polynomial type configuration and its power gives the upper bound for all the available tests. The power attained is seen in the last line of each situation. The extended max t tests are seen to keep relatively high powers in the wide range of the ordered alternatives as compared with the maximin linear tests. The linear tests with polynomial type coefficients look very good when the changepoint is located in the middle but too bad when it is in the end so that they cannot be recommended without any prior information on the configuration of mean vectors. It is just like the linear trend test in case of the simple ordered hypothesis. Another advantage of the maximal contrast type tests is that they can suggest a changepoint.

7. Application: testing sigmoidicity hypothesis

We apply the sigmoidicity test to the data of table 4 in Schmoyer (1984) which are originally from an experiment performed by Dalbey & Lock (1982). We use the normal approximation for the vector of proportions of the occurrences at respective dose levels with the dispersion matrix $\mathrm{diag}\{p_i(1-p_i)/n_i\}$, where the p_i are replaced reasonably by the maximum likelihood estimator under the sigmoidicity hypothesis obtained in the paper. Zero estimate of p_0 causes no problem if we use another expression $\boldsymbol{t} = \mathrm{diag}(N_\tau^{-1/2})(Q_K^{*\prime}\Omega Q_K^*)^{-1}Q_K^{*\prime}\bar{y}$ for $\boldsymbol{t}$ in section 3. The normal approximation will be acceptable since the number of replications at the respective dose levels are ranged from 10 to 40 and also by the cumulative nature of the statistics. Now the observed maximum is obtained as $t_o = 2.847$ at $\tau = 3$ and p value is 0.0060 by the algorithm in section 5 suggesting the assumption of sigmoidicity to be acceptable. The suggested inflection point is between $x_4 = 28$ and $x_5 = 32$ and the convexity assumption will be acceptable under the point. Hirotsu & Srivastava (2000) have discussed the simultaneous upper bounds of the risks for the data and obtained those values 0.055 and 0.035 for the lowest dose level under the monotone and convexity assumptions at lower doses, respectively, improving a naïve upper bound 0.095 based on the data at the lowest dose level only.

8. Discussion

While a parametric model gives a very efficient way of the analysis of the input–output relationship, there are often cases where those parametric models do not conform well with the data and cannot be assumed as a basis of the analysis. On the other hand a naïve pointwise estimate generally gives a very irregular and unstable response curve. Therefore those shape constraints discussed in this paper will be very useful for obtaining smooth and reasonable response curve, see Schmoyer (1984), Ramgopal *et al.* (1993) and Matzkin (1994), for example. As stated in the example of section 7 the convexity property at low doses is particularly useful for a low dose extrapolation in the toxicity analysis.

As stated in the text the isotonic regression approach will be too complicated to give an exact procedure for those extended problems. Instead the extensions of the maximin linear test (Abelson & Tukey, 1963) or the cumulative chi-squared statistic (Hirotsu, 1979) might be considered. They are, however, useful as the overall trend test and cannot suggest a changepoint. In particular the power comparisons in section 6 has shown that the linear test is not useful in those extended problems considered here.

An efficient algorithm given in section 5 depends on the simultaneous transformation of the range of integration of $S_\tau, \tau = 1, \ldots, K-2$, irrespective of the values of conditioning variables and cannot be applied as it is for the Γ sequence (Hsu, 1979), for example, where the range of integration is a function of the conditioning variables. Only for the exponential distribution another efficient algorithm is obtained (Noé, 1972).

Acknowledgements

The authors thank Prof. M. Sugihara for his valuable comments which led to the efficient algorithm of section 5. Their thanks are also due to Mr Nishihara for his help in executing the calculation.

References

Abelson, R. P. & Tukey, J. W. (1963). Efficient utilization of non-numerical information in quantitative analysis: general theory and the case of the simple order. *Ann. Math. Statist.* **34**, 1347–1369.

Dalbey, W. & Lock, S. (1982). Inhalation toxicology of diesel fuel obscurant aerosol in sprague dawley rats. ORNL/TM-8867, Oak Ridge National Laboratory, Biology Division.

Hawkins, A. J. (1977). Testing a sequence of observations for a shift in location. *J. Amer. Statist. Assoc.* **72**, 180–186.

Hirotsu, C. (1979). The cumulative chi-squares method and studentized maximal contrast method for testing an ordered alternative in a one-way analysis of variance model. *Repo. of Statisti. Applica. Res. Union Japan. Scient. Engrs.* **26**, 12–21.

Hirotsu, C. (1982). Use of cumulative efficient scores for testing ordered alternatives in discrete models. *Biometrika* **69**, 567–577.

Hirotsu, C. (1986). Cumulative chi-squared statistic as a tool for testing goodness of fit. *Biometrika* **73**, 165–173.

Hirotsu, C. (1988). A class of estimable contrasts in an age-period-cohort model. *Ann. Inst. Statist. Math.* **40**, 451–465.

Hirotsu, C. (1997). Two-way change-point model and its application. *Austral. J. Statist.* **39**, 205–218.

Hirotsu, C. & Srivastava, M. S. (2000). Simultaneous confidence intervals based on one-sided max t test. *Statist. Probab. Lett.* **49**, 25–37.

Hirotsu, C., Kuriki, S. & Haytor, A. C. (1992). Multiple comparison procedures based on the maximal component of the cumulative chi-squared statistic. *Biometrika* **79**, 381–392.

Hsu, D. A. (1979). Detecting shifts of parameter in gamma sequences with applications to stock price and air traffic flow analysis. *J. Amer. Statist. Assoc.* **74**, 31–40.

Laurie, D. P. (1996). Periodizing transformations for numerical integration. *J. Comput. Appl. Math.* **66**, 337–344.

Marcus, R. (1976). The powers of some tests of the equality of normal means against an ordered alternative. *Biometrika* **63**, 177–183.

Matzkin, R. L. (1994). Restrictions of economic theory in nonparametric methods, in *Handbook of Econometrics 4* (eds R. F. Engle and D. L. MoFaaden), 2523–2558. North-Holland, Amsterdam.

Noé, M. (1972). The calculation of distributions of two-sided Kolmogorov–Smirnov type statistics. *Ann. Math. Statist.* **43**, 58–64.

Ramgopal, P., Laud, P. W. & Smith, A. F. M. (1993). Nonparametric Bayesian bioassay with prior constraints on the shape of the potency curve. *Biometrika* **80**, 489–498.

Robertson, T., Wright, F. M. & Dykstra, R. L. (1988). *Order restricted statistical inference*. Wiley, New York.

Schmoyer, R. L. (1984). Sigmoidally constrained maximum likelihood estimation in quantal bioassay. *J. Amer. Statist. Assoc.* **79**, 448–453.

Sen, A. & Srivastava, M. S. (1975). On tests for detecting change in mean. *Ann. Statist.* **3**, 98–108.

Williams, D. A. (1971). A test for difference between treatment means when several dose levels are compared with a zero dose control. *Biometrics* **27**, 103–119.

Worsley, K. J. (1986). Confidence regions and tests for a change point in a sequence of exponential family of random variables. *Biometrika* **73**, 91–104.

Received August 2000, in final form April 2001

Chihiro Hirotsu, Faculty of Science and Technology, Meisei University, 2-1-1 Hodokubo, Hino-City, Tokyo 191-8506, Japan.
E-mail: hirotsu@ge.meisei-u.ac.jp

Appendix: corner vector of the H_3

It is easy to see that the relation (8) can be written in the matrix notation as

$$Q_K^{*\prime}\boldsymbol{\mu} \geq \mathbf{0}, \tag{16}$$

where the τth row $Q_K^{*\prime}(\tau, j),\ j = 1, \ldots, K$, of $Q_K^{*\prime}$ has only four elements not equal to zero,

$$Q_K^{*\prime}(\tau, j) =
\begin{cases}
\dfrac{1}{(x_\tau - x_{\tau+1})(x_\tau - x_{\tau+2})}, \quad j = \tau, \\
-\dfrac{1}{(x_\tau - x_{\tau+1})(x_\tau - x_{\tau+2})} - \dfrac{1}{(x_\tau - x_{\tau+2})(x_{\tau+1} - x_{\tau+2})} \\
-\dfrac{1}{(x_{\tau+1} - x_{\tau+2})(x_{\tau+1} - x_{\tau+3})}, \quad j = \tau + 1, \\
\dfrac{1}{(x_\tau - x_{\tau+2})(x_{\tau+1} - x_{\tau+2})} + \dfrac{1}{(x_{\tau+1} - x_{\tau+2})(x_{\tau+1} - x_{\tau+3})} \\
+\dfrac{1}{(x_{\tau+1} - x_{\tau+3})(x_{\tau+2} - x_{\tau+3})}, \quad j = \tau + 2, \\
-\dfrac{1}{(x_{\tau+1} - x_{\tau+3})(x_{\tau+2} - x_{\tau+3})}, \quad j = \tau + 3.
\end{cases}$$

Then it is only a tedious but not difficult task to verify $C'Q_K^* = 0$ and we can impose the restriction

$$C'\boldsymbol{\mu} = \mathbf{0} \tag{17}$$

without violating the relation (8) so that all the $\boldsymbol{\mu}$ satisfying (16) and (17) can be expressed as

$$\boldsymbol{\mu} = C(\eta_0, \eta_1, \eta_2) + Q_K^*\left(Q_K^{*\prime}Q_K^*\right)^{-1}\boldsymbol{h}$$

with $\boldsymbol{h} \geq \mathbf{0}$ and η_0, η_1, η_2 arbitrary regression coefficients.

Again it is very easy to verify

$$Q_K^{*\prime}(\boldsymbol{c}_1 \cdots \boldsymbol{c}_{K-3}) = I_{K-3}$$

and this, together with the relation $C'\boldsymbol{c}_\tau = \mathbf{0}$, implies the equality

$$Q_K^*\left(Q_K^{*\prime}Q_K^*\right)^{-1} = (\boldsymbol{c}_1 \cdots \boldsymbol{c}_{K-3})$$

showing that every model $K_3(\tau)$ indexed by τ forms $K - 3$ corners of the polyhedral cone defined by H_3.