Toutenburg, Fieger, Schaffrin:

Approximate Confidence Regions for Minimax-Linear Estimators

Sonderforschungsbereich 386, Paper 166 (1999)

Approximate Confidence Regions for Minimax-Linear Estimators

H. Toutenburg* A Fieger* B. Schaffrin**

1st September 1999

Abstract

Minimax estimation is based on the idea, that the quadratic risk function for the estimate $\hat{\beta}$ is not minimized over the entire parameter space $I\!R^K$, but only over an area $B(\beta)$ that is restricted by a priori knowledge. If all restrictions define a convex area, this area can often be enclosed in an ellipsoid of the form $B(\beta) = \{\beta : \beta' T \beta \leq r\}$. The ellipsoid has a larger volume than the cuboid. Hence, the transition to an ellipsoid as a priori information represents a weakening, but comes with an easier mathematical handling.

Deriving the linear Minimax estimator we see that it is biased and non-operationable. Using an approximation of the non-central χ^2-distribution and prior information on the variance, we get an operationable solution which is compared with OLSE with respect to the size of the corresponding confidence intervals.

1 Introduction

We consider the linear regression model

$$y = X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I) \tag{1}$$

with nonstochastic regressor matrix X of full column rank K. The sample size is T. The restriction to uncorrelated errors is not essential since it is easy to give the corresponding formulae for a covariance matrix $\sigma^2 W \neq \sigma^2 I$. If there is no further information given, the Gauß-Markov estimator for OLSE: Ordinary Least Squares Estimator

$$b = (X'X)^{-1} X' y = S^{-1} X' y \sim N(\beta, \sigma^2 S^{-1}) \tag{2}$$

with $S = X'X$ is optimal with respect to the BLUE-property. The variance factor σ^2 is estimated by

$$s^2 = (y - Xb)'(y - Xb)(T - K)^{-1} \sim \frac{\sigma^2}{T - K} \chi^2_{T-K}. \tag{3}$$

*Institut für Statistik, Ludwig-Maximilians-Universität München

**Department of Civil and Environmental Engineering and Geodetic Science, Ohio State University, 2070 Neil Ave., Columbus, OH 43210-1275, U.S.A.

Confidence Regions for β on the Basis of b: From (2) we get

$$\sigma^{-1}S^{1/2}(b-\beta) \sim N(0, I) \tag{4}$$

and thus

$$\frac{1}{K}\sigma^{-2}(b-\beta)'S(b-\beta) \sim \frac{1}{K}\chi^2_K. \tag{5}$$

As this χ^2 variable is independent of s^2, we may conclude that

$$K^{-1}s^{-2}(b-\beta)'S(b-\beta) \sim F_{K,T-K}. \tag{6}$$

From the central $F_{K,T-K}$ distribution we define the $(1-\alpha)$ fractile $F_{K,T-K}(1-\alpha)$ according to

$$P(F \leq F_{K,T-K}(1-\alpha)) = 1-\alpha. \tag{7}$$

Using these results we have

$$P\left(\frac{(b-\beta)'S(b-\beta)}{s^2K} \leq F_{K,T-K}(1-\alpha)\right) = 1-\alpha\,. \tag{8}$$

In this way we have found a simultaneous confidence region for β which is formed by the interior of the K-dimensional ellipsoid

$$\frac{1}{K}\frac{(b-\beta)'S(b-\beta)}{s^2} = F_{K,T-K}(1-\alpha). \tag{9}$$

In practice, besides the simultaneous confidence region, one might be more interested in the resulting *intervals for the components* β_i. They are determined according to Appendix A. From (75) we get the interval for the i-th component β_i $(i = 1, \dots, K)$:

$$b_i - g_i \leq \beta_i \leq b_i + g_i\,, \tag{10}$$

with

$$g_i = \sqrt{F_{K,T-K}(1-\alpha)s^2K(S^{-1})_{ii}} \tag{11}$$

where $(S^{-1})_{ii}$ is the i-th diagonal element of S^{-1} and b_i is the i-th component of b.

The length of the intervals (11) is

$$l_i = 2g_i\,. \tag{12}$$

The points of intersection of the ellipsoid (9) with the β_i-axes result from (76) as

$$\beta_i = b_i \pm \sqrt{\frac{F_{K,T-K}(1-\alpha)s^2K}{(S)_{ii}}}\,. \tag{13}$$

In the special case of a diagonal matrix $S = X'X$ (as e.g. in the case of orthogonal regressors), S^{-1} is diagonal, too, and we have $(S^{-1})_{ii} = 1/(S)_{ii}$. Hence in this case for once the points of intersection (13) with the β_i-axes coincide with the end points of the confidence intervals (10). But in general we have $(S)_{ii}^{-1} \leq (S^{-1})_{ii}$.

2 Minimax-Linear Estimation

Under the additional condition

$$\beta' B \beta \leq r \tag{14}$$

with a positive definite $(K \times K)$ matrix B and a constant $r \geq 0$, the minimax-linear estimation (MMLE) is of the form

$$b^* = (r^{-1}\sigma^2 B + S)^{-1} X' y = D^{-1} X' y \tag{15}$$

where $D = (I^{-1}\sigma^2 B + S)$; cf.,e. g. ,Rao and Toutenburg(1995, Theorem 3.9). This estimator is biased:

$$\lambda = \mathrm{bias}(b^*, \beta) = \mathrm{E}(b^*) - \beta = (D^{-1}S - I)\beta = -r^{-1}\sigma^2 D^{-1} I. \tag{16}$$

The covariance matrix $\sigma^2\,\mathrm{V}$ is

$$\sigma^2\,\mathrm{V} = \mathrm{E}\left[(b^* - \mathrm{E}(b^*))(b^* - \mathrm{E}(b^*))'\right] = \sigma^2 D^{-1} S D^{-1}. \tag{17}$$

Assuming normal distribution of ϵ thus leads to

$$\begin{aligned} b^* - \beta &\sim N(\lambda, \sigma^2\,\mathrm{V}), & (18)\\ \mathrm{V}^{-1/2}(b - \beta) &\sim N(\mathrm{V}^{-1/2}\,\lambda, \sigma^2 I), & (19)\\ \sigma^{-2}(b^* - \beta)' V^{-1} (b^* - \beta) &\sim \chi^2_K(\delta) & (20) \end{aligned}$$

with the noncentrality parameter

$$\delta = \sigma^{-2}\lambda'\,\mathrm{V}^{-1}\,\lambda. \tag{21}$$

As the MMLE b^* (15) is dependent on the unknown parameter σ^2, this estimator is not operational.

Substitution of σ^2: We confine ourselves on the substitution to σ^2 by a positive constant c and, therefore obtain the corrected estimator

$$b^*_c = D_c^{-1} X' y \tag{22}$$

with

$$\begin{aligned} D_c &= (r^{-1} c B + S), & (23)\\ \mathrm{bias}(b^*_c, \beta) &= (D_c^{-1} S - I)\beta = -r^{-1} c D_c^{-1} B \beta = \lambda_c, & (24)\\ \sigma^2\,\mathrm{V}_c &= \sigma^2 D_c^{-1} S D_c^{-1} & (25) \end{aligned}$$

and

$$\sigma^{-2}(b^*_c - \beta)'\,\mathrm{V}_c^{-1}(b^*_c - \beta) \sim \chi^2_K(\delta_c)\,, \tag{26}$$

where the noncentrality parameter δ_c is given by

$$\begin{aligned} \delta_c &= \sigma^{-2}\lambda_c'\,\mathrm{V}_c^{-1}\,\lambda_c = \sigma^{-2}\beta'(S D_c^{-1} - I) D_c S^{-1} D_c (D_c^{-1} S - I)\beta & (27)\\ &= \sigma^{-2}\beta'(S - D_c) S^{-1} (S - D_c)\beta = (\sigma^2 S^2)^{-1} c^2 \beta' B' S^{-1} B \beta\,. & (28) \end{aligned}$$

We note that δ_c is unknown, too, along with the unknown $\sigma^{-1}\beta$.

The choice of c has to be done such that the corrected MMLE b_c^* is superior to the Gauß-Markov estimator b. Based on the scalar quadratic risk of an estimator $\hat{\beta}$

$$\mathrm{R}(\hat{\beta}, a) = a' \,\mathrm{E}\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right] a \tag{29}$$

with a fixed $K \times 1$ vector $a \neq 0$, it holds that

$$\mathrm{R}(b, a) \geq \sup_c \{\mathrm{R}(b_c^*, a) : \beta' B \beta \leq r\} \,, \tag{30}$$

if (see Toutenburg, 1982, p.96)

$$c \leq 2\sigma^2 \,. \tag{31}$$

This (sufficient) condition follows from a general lemma on the robustness of the MMLE against misspecification of the additional restriction $\beta' B \beta \leq r$ since the substitution of σ^2 by c may be interpreted as a misspecified ellipsoid of the shape

$$\beta' B \beta \leq r \sigma^2 c^{-1}. \tag{32}$$

The condition (31) is practical, if a lower bound for σ^2 is known:

$$\sigma_1^2 < \sigma^2, \tag{33}$$

resulting in the choice

$$c = 2\sigma_1^2 \tag{34}$$

for c. Such a lower bound may be reclaimed using the estimator s^2 of σ^2:

$$P\left(\frac{s^2(T-K)}{\chi^2_{T-K}(1-\alpha)} \leq \sigma^2\right) = 1 - \alpha \,. \tag{35}$$

Hence one may choose $\sigma_1^2 \leq s^2(T-K)/\chi^2_{1-\alpha}$ at a $1-\alpha$ level of significance. The estimator b_c^* with $c = 2\sigma_1^2$ is called two-stage minimax linear estimator (2SMMLE).

3 Approximation of the noncentral χ^2 distribution

By formula (24.21), page 245, in Kendall and Stuart (1977), a noncentral χ^2 distribution may be approximated by a central χ^2 distribution according to

$$\chi^2_K(\delta_c) \approx a\chi^2_d \tag{36}$$

with

$$a = \frac{K + 2\delta_c}{K + \delta_c}, \quad d = \frac{(K + \delta_c)^2}{K + 2\delta_c}, \tag{37}$$

where, due to the unknown δ_c, the factor a and the number of degrees of freedom d are unknown, too.

With the approximation (36), formula (26) becomes

$$a^{-1}\sigma^{-2}(b_c^* - \beta)' \, \mathrm{V}_c^{-1}(b_c^* - \beta) \sim \chi_d^2, \tag{38}$$

i. e. approximately (in case of independence of s^2) we have

$$\frac{(b_c^* - \beta)' \, \mathrm{V}_c^{-1}(b_c^* - \beta)}{ads^2} \sim F_{d,T-K}\,. \tag{39}$$

The wanted confidence region for β at the level $1-\alpha$ is defined by the interior of the ellipsoid

$$(b_c^* - \beta)' \, \mathrm{V}_c^{-1}(b_c^* - \beta) < ads^2 F_{d,T-K}(1-\alpha). \tag{40}$$

Because of the unknown δ_c, a and d relation (40) cannot be applied directly, but only via an approximation such as the following.

Bounds for δ_c: We rewrite the noncentrality parameter δ_c (27) as follows. From

$$\mathrm{bias}(b_c^*, \beta) = \lambda_c = (D_c^{-1}S - I)\beta = -r^{-1}cD_c^{-1}B\beta, \tag{41}$$

we get

$$\begin{aligned} \delta_c &= \sigma^{-2}\lambda_c' \, \mathrm{V}_c^{-1} \, \lambda_c \\ &= \sigma^{-2}r^{-2}c^2\beta' BD_c^{-1}D_cS^{-1}D_cD_c^{-1}B\beta \\ &= \sigma^{-2}r^{-2}c^2\beta' BS^{-1}B\beta\,. \end{aligned} \tag{42}$$

Let $\lambda_{\min}(A)$ denote the minimal and $\lambda_{\max}(A)$ the maximal eigenvalue of a matrix A, respectively. Then it is well-known that "Raleigh's inequalities"

$$0 \le \beta' B\beta\lambda_{\min}(B^{1/2}S^{-1}B^{1/2}) \le \beta' BS^{-1}B\beta \le r\lambda_{\max}(B^{1/2}S^{-1}B^{1/2}) \tag{43}$$

hold true, yielding for a general c and with the inequality (33) at first

$$\delta_c \le \sigma_1^{-2}r^{-1}c^2\lambda_{\max}(B^{1/2}S^{-1}B^{1/2}), \tag{44}$$

and for $c = 2\sigma_1^2$ especially

$$\delta_c \le 2cr^{-1}\lambda_{\max}(B^{1/2}S^{-1}B^{1/2}) = \delta_{\mathrm{o}}. \tag{45}$$

Hence, the upper bound δ_0 for δ_c can be calculated for any c.

Using this inequality, we get for the coefficients a and d of the approximation (36)

$$a \le \frac{K + 2\delta_{\mathrm{o}}}{K + \delta_{\mathrm{o}}} = a_{\mathrm{o}} \tag{46}$$

and

$$d \le \frac{(K + \delta_{\mathrm{o}})^2}{K + 2\delta_{\mathrm{o}}} = d_{\mathrm{o}}, \tag{47}$$

i.e.

$$ad = K + \delta_c \leq K + \delta_o = a_o d_o. \tag{48}$$

The approximate confidence region for β then becomes

$$\{\beta : (b_c^* - \beta)' \, \mathrm{V}_c^{-1} (b_c^* - \beta) < (K + \delta_o) s^2 F_{d_o, T-K}(1-\alpha)\}. \tag{49}$$

We have $(K + \delta_c) \leq (K + \delta_o)$, but $F_{d,T-K}(1-\alpha) \geq F_{d_o,T-K}(1-\alpha)$ for realistic choices of α and $T - K \geq 3$. Thus the impact of changing the actual parameter to its maximal value δ_o, on the volume of the confidence region (49) used in practice, has to be analysed numerically. Simulations (see Section 4) were carried out which show that using δ_o instead of δ_c will increase the volume of the confidence region.

With the abbreviation

$$g_i^o = \sqrt{F_{d_o,T-K}(1-\alpha) s^2 (K + \delta_o)(\mathrm{V}_c)_{ii}}\ , \tag{50}$$

it follows from (49), that the confidence intervals for the components from β may be written as

$$\mathrm{KI}_i = [b_{c_i}^* - g_i^o < \beta_i < b_{c_i}^* + g_i^o]\,. \tag{51}$$

4 Properties of Efficiency

Let us now investigate the efficiency of the proposed solution. Assume that the confidence level $1 - \alpha$ is fixed. Replacing δ_c by the least favourable value δ_o has influence on the length of the confidence intervals.

a) True, but unknown confidence region (40) on the basis of δ_c. Length of the confidence interval:

$$2g_i^c = 2\sqrt{F_{d,T-K}(1-\alpha) s^2 (K + \delta_c)(\mathrm{V}_c)_{ii}}.$$

b) Practical confidence region (49) on the base of δ_o. Length of the confidence interval according to (50):

$$2g_i^o = 2\sqrt{F_{d_o,T-K}(1-\alpha) s^2 (K + \delta_o)(\mathrm{V}_c)_{ii}}.$$

By defining the ratio

$$\frac{\text{Length of the interval on the basis of } \delta_o}{\text{Length of the interval on the basis of } \delta_c}$$

we get (for all $i = 1, \ldots, K$) the same stretching factor

$$f = f(\delta_c, \delta_o, K, T-K) = \frac{2g_i^o}{2g_i^c} = \sqrt{\frac{(K+\delta_o) F_{d_o,T-K}(1-\alpha)}{(K+\delta_c) F_{d,T-K}(1-\alpha)}}\,. \tag{52}$$

For given values of δ_c where ($\delta_c = 0.1$ and $\delta_c = 1$) and for $T - K = 10$ and $T - K = 33 - K$, respectively, we have calculated the stretching factor in dependence of δ_o and varying values of K (Figures 3, 4 and 5). The stretching factor is decreasing with increasing K (number of regressors) and is increasing with the distance $(\delta_o - \delta_c)$; see Fig. 3–4.

Another kind of rating the quality of the practical confidence region (49) is to determine the equivalent confidence level $1 - \alpha_1$ of the true (but unknown) confidence region (40).

The true confidence region is defined approximately through

$$P\left(\frac{(b_c^* - \beta)' \, \mathrm{V}_c^{-1}(b_c^* - \beta)}{ads^2} \leq F_{d,T-K}(1-\alpha)\right) = 1 - \alpha. \tag{53}$$

Due to the replacement of δ_c by its maximum δ_o, we instead determine an increased confidence ellipsoid by

$$P\left(\frac{(b_c^* - \beta)' \, \mathrm{V}_c^{-1}(b_c^* - \beta)}{a_o d_o s^2} \leq F_{d_o,T-K}(1-\alpha)\right) = 1 - \alpha_1. \tag{54}$$

Hence, by combination of (53) and (54), we find for the true (and smaller) confidence region

$$P\left(\frac{(b_c^* - \beta)' \, \mathrm{V}_c^{-1}(b_c' - \beta)}{ads^2} \leq f^2 F_{d,T-K}(1-\alpha)\right) = 1 - \alpha_1 \geq 1 - \alpha \tag{55}$$

with $f \geq 1$ from (52). Replacing the unknown noncentrality parameter δ_c by its maximum δ_o results in an increase of the confidence level, as we have $\alpha_1 \leq \alpha$ (Figures 5–6 present values of α_1 for varying values of T and K).

As a consequence in practice we choose a smaller confidence level of e. g. $1 - \alpha = 0.90$ to reach a real confidence level of $1 - \alpha_1 < 1$ (also for greater distances $\delta_o - \delta_c$).

The stretching factor f *and,* moreover, the increase of the confidence level are increasing with $\delta_o - \delta_c$. For model and data given the distance $\delta_o - \delta_c$ may be approximately determined by

$$\delta_o - \delta_c \leq \delta_o - \delta_u \leq \delta_o \tag{56}$$

where, according to (42) and (43),

$$\delta_u = 4\sigma_1^2 r^{-2} \left[\, \lambda_{\min}(B^{1/2} S^{-1} B^{1/2})\right] \beta' B \beta \geq 0 \tag{57}$$

turns out to be a lower bound of the true noncentrality parameter δ_c. The upper bound δ_o is calculated for concrete models such that it becomes possible to estimate the maximum stretch factor f and the maximal increase of the confidence level from $1-\alpha$ to $1-\alpha_1$. In this way the practicality of the proposed method is given in addition to the estimation of its efficiency.

If the ellipsoid of the prior information is not centred in the origin but in a general mid point vector $\beta_0 \neq 0$, i. e.

$$(\beta - \beta_0)' B (\beta - \beta_0) \leq r, \tag{58}$$

then the MMLE becomes

$$b^*(\beta_0) = \beta_0 + D^{-1} X'(y - X\beta_0) \tag{59}$$

with

$$\text{bias}(b^*(\beta_0), \beta) = (D^{-1}S - I)(\beta - \beta_0) \tag{60}$$

and (see (17))

$$\text{V}(b^*(\beta_0)) = \text{V}(b^*) = V\,. \tag{61}$$

All the preceding results remain valid if we replace for λ in (16) and δ_c in (24) the vector β by $(\beta - \beta_0)$, provided that δ in (21) and δ_c in (27) is defined with the accordingly changed λ and λ_c.

5 Comparing the Volumes

The definition of a confidence ellipsoid is based on the assumption that the unknown parameter β is contained with probability $1 - \alpha$ in the random ellipsoid. If one has the choice between alternative ellipsoids, one would choose the ellipsoid with the smallest volume. In other words, the MDE-superiority of the MMLE with respect to the Gauß-Markov estimator in the sense of (14) does not necessarily lead to a preference of the ellipsoids based on the MMLE. Hence in the following we determine the volume of both ellipsoids. The volume of the T-dimensional unit sphere

$$x'x \leq 1$$

(x being a $T \times 1$-Vector) is given as

$$\text{Vol}_\text{E} = \frac{\pi^{T/2}}{\Gamma(1 + \frac{T}{2})}. \tag{62}$$

For an ellipsoid $x'Ax \leq 1$ with a positive definite matrix A, the volume is

$$\text{Vol}(A) = \text{Vol}_\text{E}|A|^{-1/2}. \tag{63}$$

a) Gauß-Markov Estimator The confidence ellipsoid for β on the basis of the Gauß-Markov estimator b in (2) is, according to (8),

$$\frac{1}{s^2 K F_{K,T-K}(1-\alpha)}(b - \beta)'S(b - \beta) \leq 1\,, \tag{64}$$

thus the volume becomes

$$\text{Vol}(b) = (s^2 K F_{K,T-K}(1-\alpha))^{K/2}|S|^{-1/2}\text{Vol}_\text{E}. \tag{65}$$

b) MMLE Based on the approximations (36) and (48), the confidence region using the MMLE b_c^* was (see (49))

$$\frac{1}{(K + \delta_o)s^2 F_{d_o,T-K}(1-\alpha)}(b_c^* - \beta)'\text{V}_c^{-1}(b_c^* - \beta) \leq 1, \tag{66}$$

hence its volume is

$$\text{Vol}(b_c^*) = ((K + \delta_o)s^2 F_{d_o,T-K}(1-\alpha))^{K/2}|\,\text{V}_c^{-1}\,|^{-1/2}\text{Vol}_\text{E}\,. \tag{67}$$

Comparing both volumes with each other gives

$$q = \frac{\mathrm{Vol}(b_c^*)}{\mathrm{Vol}(b)} = [f(0, \delta_o, K, T-K)]^K \frac{|X'X|^{1/2}}{|\,\mathrm{V}_c^{-1}\,|^{1/2}} \tag{68}$$

where $f(0, \delta_o, K, T-K)$ is the maximal stretch factor (52) for the lower bound $\delta_u = 0$ of the noncentrality parameters δ_c. $\delta_u = 0$ corresponds to $T \to \infty$,i. e. to the change from the MMLE to the Gauß-Markov estimator. The MMLE b_c has smaller variance than the Gauß-Markov estimator b:

$$0 \leq (X'X)^{-1} - \mathrm{V}_c \quad \text{(nonnegative definite,)}$$

i. e. we have

$$0 \leq \mathrm{V}_c^{-1} - X'X = 2r^{-1}cB + r^{-1}c^2 BSB = C$$

or

$$\mathrm{V}_c^{-1} = X'X + C \text{ with } C \geq 0 \text{ (nonnegative definite).} \tag{69}$$

From (69) we may conclude that

$$|X'X| \leq |\,\mathrm{V}_c^{-1}\,|\,,$$

and thus

$$|I + r^{-1}cBS^{-1}|^{-1} = \frac{|X'X|^{1/2}}{|\,\mathrm{V}_c^{-1}\,|^{1/2}} \leq 1\,. \tag{70}$$

So the relation (68) between both volumes turns out to be the product of a function $f \geq 1$ and the expression (70) which is ≤ 1.

The ratio (68) has to be investigated for a concrete model and given data, as δ_o (and hence f) and the quantity (70) are dependent on the data as well as on the strength of the additional condition.

Let $X'X = S$ and assume the condition $\beta' S \beta \leq r$. Then according to Section 3 we have

$$\begin{aligned} \mathrm{V}_c^{-1} &= (r^{-2}c^2 + 1 + 2r^{-1}c)S \\ &= (r^{-1}c + 1)^2 S \end{aligned}$$

and

$$|\,\mathrm{V}_c^{-1}\,| = (r^{-1}c + 1)^{2K} |S|. \tag{71}$$

Analogously, from (45) with $c = 2\sigma_1^2$, we get

$$\delta_o = 2r^{-1}c\,. \tag{72}$$

This results in a change of the relation of the volumes (68) to

$$q = q(r^{-1}c) = \left(\frac{f(0, 2r^{-1}c, K, T-K)}{(r^{-1}c + 1)} \right)^K . \tag{73}$$

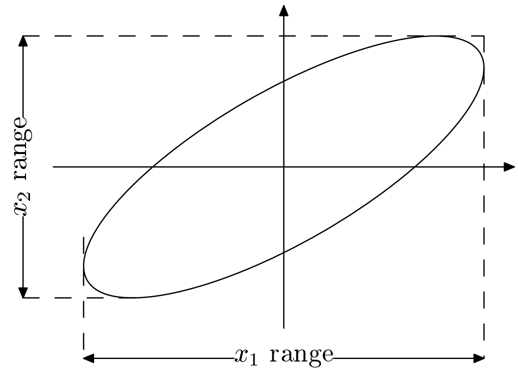

Figure 1: Region of the components x_1 and x_2 of an ellipsoid $x'Ax = r$

A Appendix

Assume an ellipsoid

$$x'Ax = r$$

with the positive definite matrix A and the $1 \times T$ vector $x' = (x_1, \ldots, x_T)$.

We determine the regions of the x_i components for the ellipsoid. Without loss of generality, we solve this problem for the first component x_1 only, this is equivalent to finding an extremum under linear constraints.

Let $e_1' = (1, 0, \ldots, 0)$ and μ be a Lagrange muliplier. Further, let

$$\begin{aligned} f(x) &= x_1 = e_1'x \,, \\ g(x) &= x'Ax - r \end{aligned}$$

and

$$F(x) = f(x) + \mu g(x).$$

Then we have to solve

$$F(x) = \text{stationary}_{x,y} \,,$$

which leads to the necessary normal equations

$$\begin{aligned} \frac{\partial F(x)}{\partial x} &= e_1 + 2\hat{\mu}Ax = 0, \qquad (74) \\ \frac{\partial F(x)}{\partial \mu} &= x'Ax - r = 0. \end{aligned}$$

From (74) it follows that:

$$x'e_1 + 2\hat{\mu}x'Ax = 0 \,,$$

thus we get

$$2\hat{\mu} = -\frac{x_1}{r} \,.$$

Inserting this into (74) gives

$$\begin{aligned} e_1 + 2\hat{\mu}Ax &= e_1 - \frac{x_1}{r}Ax = 0\,, \\ x &= A^{-1}e_1\frac{r}{x_1} \end{aligned}$$

and therefore

$$e_1'x = x_1 = e_1'A^{-1}e_1\frac{r}{x_1},$$

or

$$x_1 = \pm\sqrt{r(A^{-1})_{11}},$$

with $(A^{-1})_{11}$ as the first diagonal element of the matrix A^{-1}. In case that the ellipsoid is not centered in the origin:

$$(x - x_0)'A(x - x_0) = r$$

the regions of the x_i components become

$$x_{0_i} - \sqrt{r(A^{-1})_{ii}} \leq x_i \leq x_{0_i} + \sqrt{r(A^{-1})_{ii}}\,. \tag{75}$$

The intersection points of the ellipsoid $x'Ax = r$ with the coordinate axes follow from

$$(0,\dots,x_t - x_{0i}, 0, \dots, 0)A(0,\dots,x_t - x_{0i}, 0,\dots,0)' = (x_t - x_{0i})^2(A)_{ii} = r\,,$$

as

$$x_i = x_{0i} \pm \sqrt{\frac{r}{(A)_{ii}}}. \tag{76}$$

References

Kendall, M. G. and Stuart, A. (1977). *The Advanced Theory of Statistics*, Griffin.

Rao, C. R. and Toutenburg, H. (1995). *Linear Models: Least Squares and Alternatives*, Springer.

Toutenburg, H. (1982). *Prior Information in Linear Models*, Wiley.

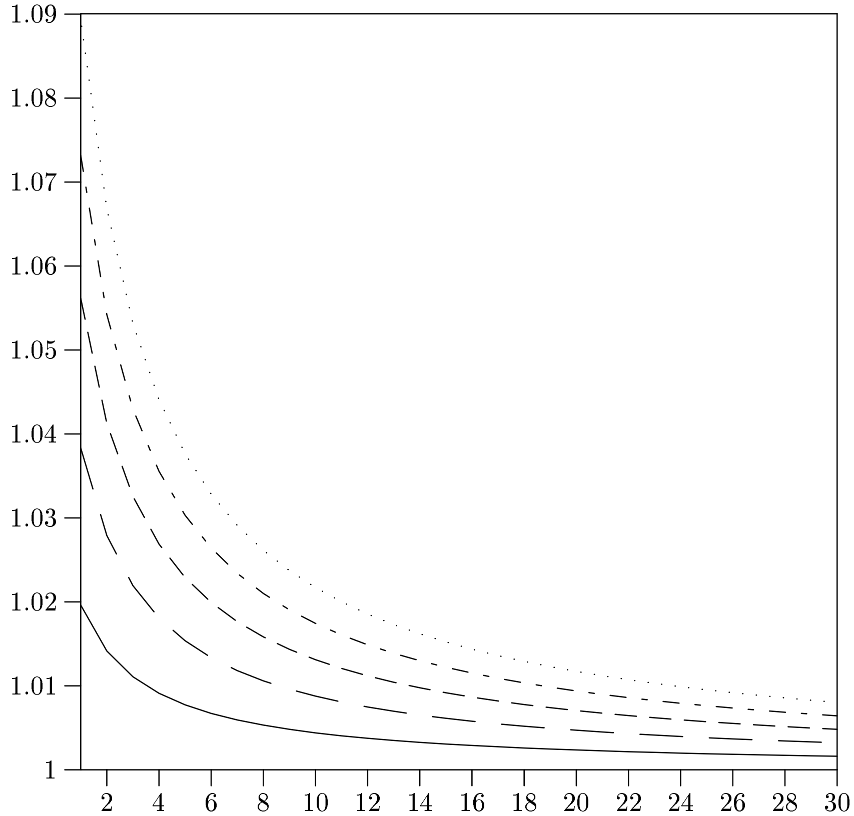

Figure 2: Stretching factor f (vertical axis) depending on K (horizontal axis). With increasing K the stretching factor is decreasing. Results are presented for $\delta_c = 1$, $T = 33$ and additionally varying δ_o starting from $\delta_c + 0.1$ (solid line, step 0.1). With increasing difference $(\delta_o - \delta_c)$ the stretching factor increases; see also Figure 4

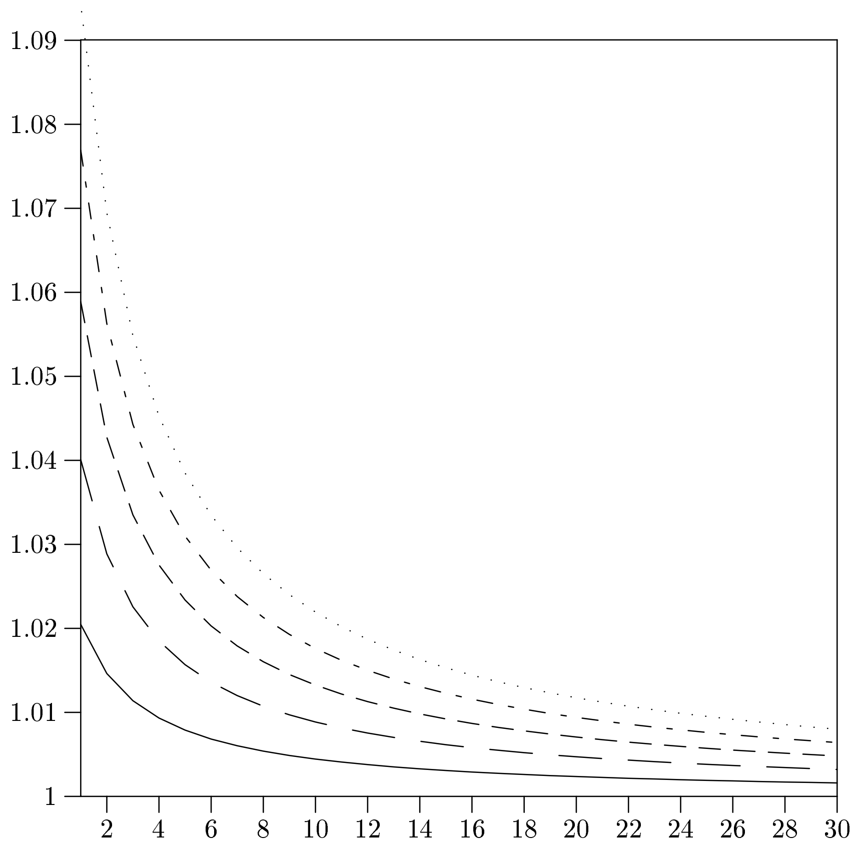

Figure 3: Stretching factor f (vertical axis) depending on K (horizontal axis). With increasing K the stretching factor is decreasing. Results are presented for $\delta_c = 1$, $T = K + 10$ and additionally varying δ_o from $\delta_c + 0.1$ (solid line, step 0.1). With increasing difference $(\delta_o - \delta_c)$ the stretching factor increases; see also Figure 4

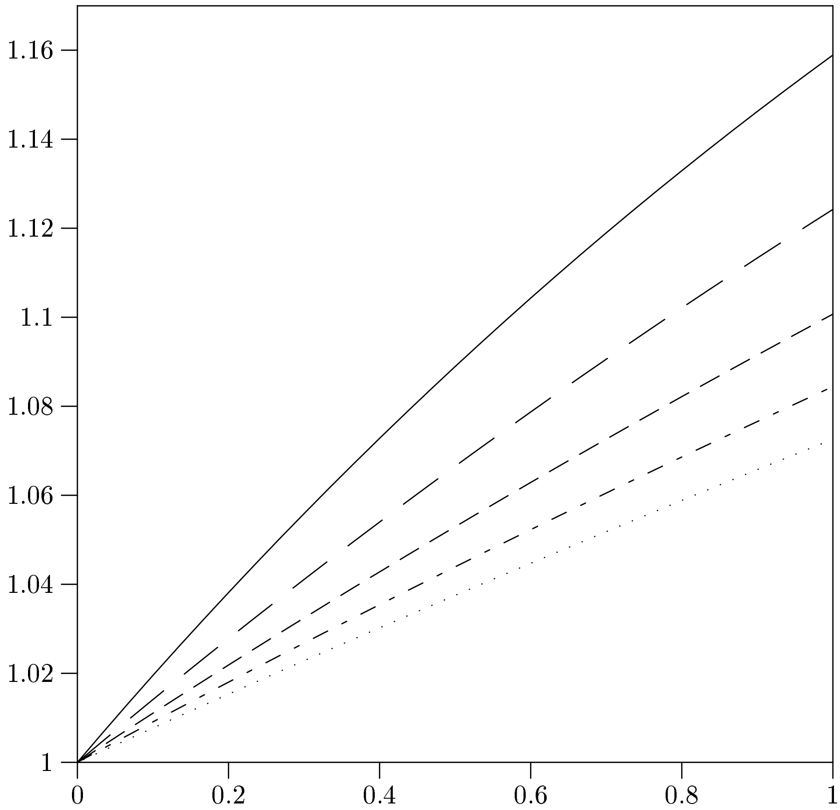

Figure 4: Stretching factor f (vertical axis) depending on the difference $(\delta_o - \delta_c)$ (horizontal axis). With increasing difference $(\delta_o - \delta_c)$ the stretching factor increases. Results are presented for $\delta_c = 1$, $T = 33$ and additionally varying K from 1 (solid line) to 5. With increasing K the stretching factor decreases; see also Figure 2

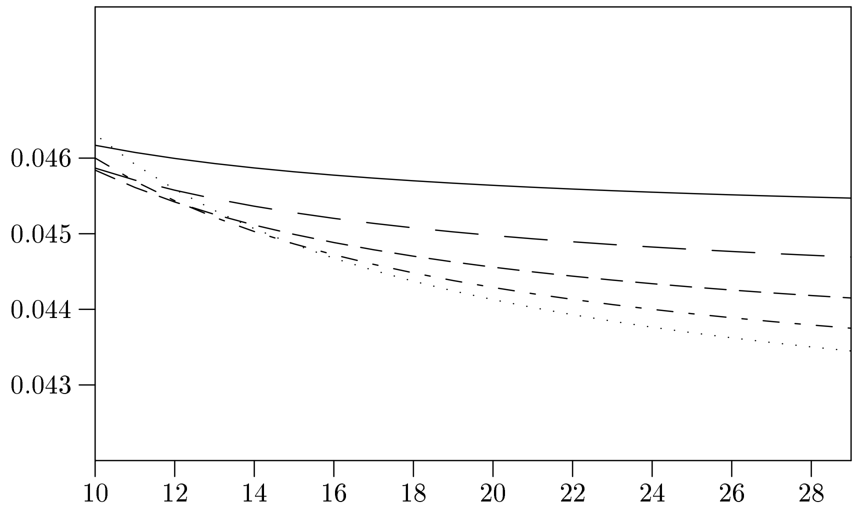

Figure 5: Confidence level α_1 (vertical axis) depending on T (horizontal axis) for $f^2 = 1.02^2$ and $\alpha = 0.05$. Additionally varying K from 1 (solid line) to 5. α_1 decreases with increasing K.

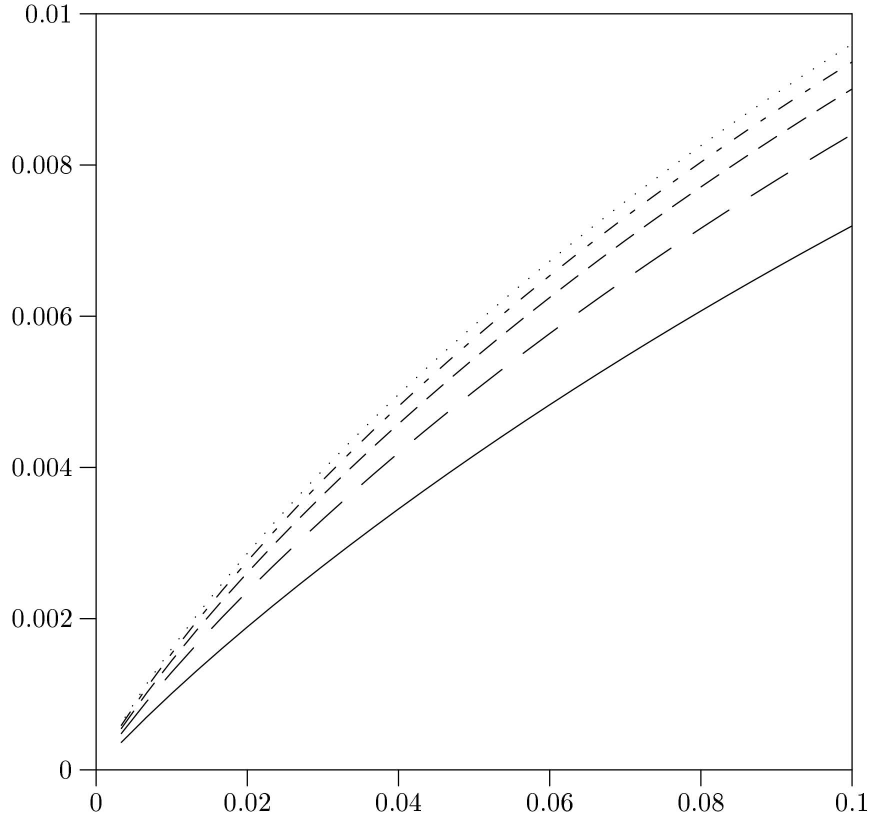

Figure 6: Difference of the confidence level $(\alpha - \alpha_1)$ (vertical axis) depending on α (horizontal axis) for $f^2 = 1.02^2$ and $K = 3$. Additionally varying $T = 10$ (solid line), $15, 20, 25, 30$. With increasing T the difference $\alpha - \alpha_1$ increases.