ELSEVIER

Statistics & Probability Letters 44 (1999) 343–350

STATISTICS & PROBABILITY LETTERS

www.elsevier.nl/locate/stapro

Robustness properties of dispersion estimators

Marc G. Genton *, Yanyuan Ma

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachussetts Avenue, Cambridge, MA 02139-4307, USA

Received June 1998; received in revised form December 1998

Abstract

In this paper, we derive the influence function of dispersion estimators, based on a scale approach. The relation between the gross-error sensitivity of dispersion estimators and the one of the underlying scale estimator is described. We show that for the bivariate Gaussian distributions, the asymptotic variance of covariance estimators is minimal in the independent case, and is strictly increasing with the absolute value of the underlying covariance. The behavior of the asymptotic variance of correlation estimators seems to be the opposite, i.e. maximal for independent data, and strictly decreasing with the absolute value of the underlying correlation. In particular, dispersion estimators based on M-estimators of scale are studied closely. The one based on the median absolute deviation is the most B-robust in the class of symmetric estimators. Some other examples are analyzed, based on the maximum likelihood and the Welsch estimator of scale.

Keywords: Covariance; Correlation; Influence function; Asymptotic variance; M-estimator; B-robust; Most B-robust

1. Introduction

The dispersion between two random variables, i.e. the covariance or the correlation, is a quantity of great interest in statistics, since it provides a measure of association or interdependence between two characteristics. Regression is a typical example where such quantities are used. In particular for several random variables, dispersion matrices are the cornerstones of many multivariate techniques (e.g. Mardia et al., 1979). Therefore, reliable dispersion estimators are of prime importance. Unfortunately, classical sample dispersion estimators are known to be very sensitive to outlying values in the data, due to gross errors, measurement mistakes, faulty recording. In this paper, we address the issue of the robustness of dispersion estimators, by means of the influence function.

Traditionally, estimation of the covariance θ between two random variables X and Y is based on a location approach, since

$$\mathrm{Cov}(X,Y) = E[(X - E(X))(Y - E(Y))], \tag{1}$$

* Corresponding author.
E-mail addresses: genton@math.mit.edu (M.G. Genton), yanyuan@math.mit.edu (Y. Ma)

0167-7152/99/$ - see front matter
PII: S0167-7152(99)00025-5

yielding for example the unbiased sample estimator:

$$\hat{\theta} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}), \tag{2}$$

where $\{X_1, Y_1\}, \ldots, \{X_n, Y_n\}$ is a sample of size n, $\overline{X} = (1/n)\sum_{i=1}^{n} X_i$, and $\overline{Y} = (1/n)\sum_{i=1}^{n} Y_i$. However, covariance estimation can also be based on a scale approach, by means of the following identity (Huber, 1981; Gnanadesikan, 1997):

$$\mathrm{Cov}(X, Y) = \frac{1}{4\alpha\beta}[\mathrm{Var}(\alpha X + \beta Y) - \mathrm{Var}(\alpha X - \beta Y)], \quad \forall \alpha, \beta \in \mathbb{R}. \tag{3}$$

The choice of a robust estimator of the variance in (3) produces a robust estimator of the covariance θ between X and Y. Similarly, estimation of the correlation ρ between X and Y can be based on a scale approach, by means of

$$\mathrm{Corr}(X, Y) = \frac{1}{4\alpha\beta\sigma_X\sigma_Y}[\mathrm{Var}(\alpha X + \beta Y) - \mathrm{Var}(\alpha X - \beta Y)], \quad \forall \alpha, \beta \in \mathbb{R}, \tag{4}$$

where $\sigma_X = \sqrt{\mathrm{Var}(X)}$ and $\sigma_Y = \sqrt{\mathrm{Var}(Y)}$. Here again, the choice of a robust estimator of the variance in (4) produces a robust estimator of the correlation ρ between X and Y. In general, X and Y may be measured in different units, and the choice $\alpha = 1/\sigma_X$ and $\beta = 1/\sigma_Y$ is recommended (Gnanadesikan and Kettenring, 1972) in Eqs. (3) and (4). The main advantage of the scale approach is that the influence functions of both covariance and correlation estimators will depend on the influence function of the scale estimator only. On the contrary, with a location approach, it would depend on the influence function of scale and location estimators. Moreover, the distribution of linear combinations of random variables as in (3) and (4) is much easier to handle than the distribution of products of random variables as in (1).

2. The influence function of dispersion estimators

Consider a sample $Z_1, \ldots, Z_n$ and a scale estimator $S_n(Z_1, \ldots, Z_n)$, i.e. such that $S_n(aZ_1 + b, \ldots, aZ_n + b) = |a|S_n(Z_1, \ldots, Z_n)$, $\forall a, b \in \mathbb{R}$. We write $S_n(Z_1, \ldots, Z_n) = S_n(F_n)$, where $F_n(z) = (1/n)\sum_{i=1}^{n} \Delta_{Z_i}(z)$ is the empirical distribution, and Δ_{Z_i} is the Dirac function with jump at Z_i. Let $S(F)$ be the corresponding statistical functional of scale such that $S(F_n) = S_n(F_n)$. The influence function (Hampel, 1974) of S at a distribution F is defined by

$$\mathrm{IF}(u; S, F) = \lim_{\varepsilon \to 0^+} \frac{S((1-\varepsilon)F + \varepsilon\Delta_u) - S(F)}{\varepsilon}, \tag{5}$$

in those u where this limit exists. The importance of the influence function lies in its heuristic interpretation: it describes the effect of an infinitesimal contamination at the point u on the estimate, standardized by the mass of the contamination, i.e., it measures the asymptotic bias caused by the contamination in the observations. The gross-error sensitivity (Hampel et al., 1986) defined by $\gamma^*(S, F) = \sup_u |\mathrm{IF}(u; S, F)|$, measures the worst asymptotic bias due to the contamination. If $\gamma^*(S, F) < \infty$, the estimator is said to be B-robust, i.e., robust with respect to the bias.

Let Θ be a statistical functional of covariance corresponding to a covariance estimator $\hat{\theta}$ based on Eq. (3):

$$\Theta(\boldsymbol{F}) = \frac{1}{4\alpha\beta}[S^2(F_+) - S^2(F_-)], \tag{6}$$

where $\boldsymbol{F}$ is a bivariate distribution with marginal distributions F_X and F_Y, and F_+ and F_- denote the distributions of $\alpha X + \beta Y$ and $\alpha X - \beta Y$, respectively. For simplicity, we assume that F_X and F_Y, both have a mean zero. A natural way to define the influence function of Θ is through the influence function of S. Note that the influence function describes the first-order sensitivity of the estimator to contamination, and thus has similar properties as the usual first derivative. In the following derivations, we frequently use the properties that $\mathrm{IF}(u,S,F)=\sigma\mathrm{IF}(u/\sigma,S,\tilde{F})$ and $\mathrm{IF}(u,h(S),F)=h'(S(F))\mathrm{IF}(u,S,F)$ (Hampel et al., 1986, p. 232 and 259; Ma and Genton, 1998), where $\tilde{F}$ is the standardized distribution of F, i.e. $\tilde{F}(z)=F(\sigma_Z z)$, and h is a real differentiable function. From Eq. (6), we get the following influence function for Θ:

$$\mathrm{IF}((u,v);\Theta,\boldsymbol{F})=\frac{1}{4\alpha\beta}[\mathrm{IF}(\alpha u+\beta v;S^2,F_+)-\mathrm{IF}(\alpha u-\beta v;S^2,F_-)]$$

$$=\frac{1}{4\alpha\beta}2S(F_+)\mathrm{IF}(\alpha u+\beta v;S,F_+)-\frac{1}{4\alpha\beta}2S(F_-)\mathrm{IF}(\alpha u-\beta v;S,F_-). \tag{7}$$

Defining the influence function of a bivariate estimator through the influence function of a univariate estimator, as in Eq. (7), provides a way to generalize the unidimensional Dirac function Δ_u to a bidimensional Dirac function. Note that in this definition, the perturbations we consider depend on the choice of the covariance estimator: they are, respectively, perturbations along $\alpha u+\beta v$ and $\alpha u-\beta v$ directions. In fact, this is a typical method to reduce a higher dimensional problem to a lower dimensional one. Using $\alpha=1/\sigma_X$ and $\beta=1/\sigma_Y$, Eq. (7) becomes

$$\mathrm{IF}((u,v);\Theta,\boldsymbol{F})=\frac{\sigma_X\sigma_Y}{2}\left[\mathrm{IF}\left(\left(\frac{u}{\sigma_X}+\frac{v}{\sigma_Y}\right)\Big/\sigma_+;S,\tilde{F}_+\right)\sigma_+^2-\mathrm{IF}\left(\left(\frac{u}{\sigma_X}-\frac{v}{\sigma_Y}\right)\Big/\sigma_-;S,\tilde{F}_-\right)\sigma_-^2\right], \tag{8}$$

where $\sigma_+=S(F_+)$ and $\sigma_-=S(F_-)$.

Let R be a statistical functional of correlation corresponding to a correlation estimator $\hat{\rho}$ based on Eq. (4):

$$R(\boldsymbol{F})=\frac{1}{4\alpha\beta S(F_X)S(F_Y)}[S^2(F_+)-S^2(F_-)]. \tag{9}$$

Similar to the covariance case, the influence function of R is

$$\mathrm{IF}((u,v);R,\boldsymbol{F})$$
$$=\mathrm{IF}((u,v);\Theta,\boldsymbol{F})\frac{1}{S(F_X)S(F_Y)}-\frac{\Theta(\boldsymbol{F})}{S^2(F_X)S^2(F_Y)}(\mathrm{IF}(u;S,F_X)S(F_Y)+\mathrm{IF}(v;S,F_Y)S(F_X)). \tag{10}$$

Using Eqs. (6) and (7), as well as $\alpha=1/\sigma_X$ and $\beta=1/\sigma_Y$, Eq. (10) becomes

$$\mathrm{IF}((u,v);R,\boldsymbol{F})=\frac{1}{\sigma_X\sigma_Y}\mathrm{IF}((u,v);\Theta,\boldsymbol{F})-\rho(\mathrm{IF}(u/\sigma_X;S,\tilde{F}_X)+\mathrm{IF}(v/\sigma_Y;S,\tilde{F}_Y)). \tag{11}$$

The links between the gross-error sensitivities for scale and for dispersion estimators are given in the next two propositions. Let us define $\gamma_+^*(S,F)=\sup_u \mathrm{IF}(u;S,F)$, $\gamma_-^*(S,F)=-\inf_u IF(u;S,F)$, $\gamma^*(\Theta,\boldsymbol{F})=\sup_{u,v}|\mathrm{IF}((u,v);\Theta,\boldsymbol{F})|$, $\gamma^*(R,\boldsymbol{F})=\sup_{u,v}|\mathrm{IF}((u,v);R,\boldsymbol{F})|$.

Proposition 1. *Let θ be the covariance between two random variables X and Y, and Θ be a statistical functional of covariance based on a statistical functional S of scale. The gross-error sensitivity of Θ is*

$$\gamma^*(\Theta,\boldsymbol{F})=\frac{\sigma_X\sigma_Y}{2}\max(\sigma_+^2\gamma_+^*(S,\tilde{F}_+)+\sigma_-^2\gamma_-^*(S,\tilde{F}_-),\ \sigma_+^2\gamma_-^*(S,\tilde{F}_+)+\sigma_-^2\gamma_+^*(S,\tilde{F}_-)).$$

In particular, when $\gamma_+^*(S,\tilde{F}_\pm)=\gamma_-^*(S,\tilde{F}_\pm)=\gamma^*(S,\tilde{F}_\pm)$,

$$\gamma^*(\Theta,\boldsymbol{F})=(\sigma_X\sigma_Y+\theta)\gamma^*(S,\tilde{F}_+)+(\sigma_X\sigma_Y-\theta)\gamma^*(S,\tilde{F}_-).$$

Proof. From Eq. (8), the influence function $\mathrm{IF}((u,v);\Theta,\boldsymbol{F})$ must be bounded between $-(\sigma_X\sigma_Y/2)\ (\sigma_+^2\gamma_-^*(S,\tilde{F}_+)+\sigma_-^2\gamma_+^*(S,\tilde{F}_-))$ and $(\sigma_X\sigma_Y/2)\ (\sigma_+^2\gamma_+^*(S,\tilde{F}_+)+\sigma_-^2\gamma_-^*(S,\tilde{F}_-))$. Because the supremum and infimum of the influence function of S can be reached simultaneously, i.e., at the same (u,v), the two bounds are tight. In particular, when $\gamma_+^*(S,\tilde F_\pm)=\gamma_-^*(S,\tilde F_\pm)=\gamma^*(S,\tilde F_\pm)$, the two bounds have the same absolute value $(\sigma_X\sigma_Y/2)\ (\sigma_+^2\gamma^*(S,\tilde F_+)+\sigma_-^2\gamma^*(S,\tilde F_-))=(\sigma_X\sigma_Y+\theta)\gamma^*(S,\tilde F_+)+(\sigma_X\sigma_Y-\theta)\gamma^*(S,\tilde F_-)$. □

Proposition 2. *Let ρ be the correlation between two random variables X and Y, and R be a statistical functional of correlation based on a statistical functional S of scale. The gross-error sensitivity of R is*

$$\begin{aligned}\gamma^*(R,\boldsymbol F)=\max\ &(\tfrac12\sigma_+^2\gamma_+^*(S,\tilde F_+)+\tfrac12\sigma_-^2\gamma_-^*(S,\tilde F_-)+\rho(\gamma_-^*(S,\tilde F_X)+\gamma_-^*(S,\tilde F_Y)),\\ &(\tfrac12\sigma_+^2\gamma_-^*(S,\tilde F_+)+\tfrac12\sigma_-^2\gamma_+^*(S,\tilde F_-)+\rho(\gamma_+^*(S,\tilde F_X)+\gamma_+^*(S,\tilde F_Y)))\quad \textit{for } \rho\geqslant 0;\end{aligned}$$

$$\begin{aligned}\gamma^*(R,\boldsymbol F)=\max\ &(\tfrac12\sigma_+^2\gamma_+^*(S,\tilde F_+)+\tfrac12\sigma_-^2\gamma_-^*(S,\tilde F_-)-\rho(\gamma_+^*(S,\tilde F_X)+\gamma_+^*(S,\tilde F_Y)),\\ &(\tfrac12\sigma_+^2\gamma_-^*(S,\tilde F_+)+\tfrac12\sigma_-^2\gamma_+^*(S,\tilde F_-)-\rho(\gamma_-^*(S,\tilde F_X)+\gamma_-^*(S,\tilde F_Y)))\quad \textit{for } \rho< 0.\end{aligned}$$

In particular, when $\gamma_+^*(S,\tilde F_\pm)=\gamma_-^*(S,\tilde F_\pm)=\gamma^*(S,\tilde F_\pm)$:

$$\gamma^*(R,\boldsymbol F)=\frac{\sigma_+^2}{2}\gamma^*(S,\tilde F_+)+\frac{\sigma_-^2}{2}\gamma^*(S,\tilde F_-)+|\rho|(\gamma^*(S,\tilde F_X)+\gamma^*(S,\tilde F_Y)).$$

Proof. Similar to Proposition 1. □

Propositions 1 and 2 tell us that the dispersion estimators are B-robust if the underlying scale estimators are B-robust. The most interesting M-estimators of scale satisfy $\gamma_+^*(S,\tilde F_\pm)\geqslant\gamma_-^*(S,\tilde F_\pm)$, with equality when they have 50% breakdown point (Huber, 1981; Genton and Rousseeuw, 1995). The opposite situation leads to implosion of the scale estimator, as well as to lower efficiency. Observe that often $\tilde F_+=\tilde F_-=\tilde F_X=\tilde F_Y$ in Propositions 1 and 2, yielding further simplifications. For instance, this is the case for multivariate Gaussian distributions, and even for some specific members of the more general class of elliptically contoured distributions (Fang and Zhang, 1990), like multivariate t or multivariate Cauchy distributions. Conditions for this property to hold can be found in Kano (1994). Note that in order to compare the gross-errors sensitivities of two dispersion estimators, one should standardize them (Hampel et al., 1986, pp. 228–229), for example with respect to their variances (self-standardized), or to the Fisher information (information-standardized).

3. Comparison of asymptotic variances of some dispersion estimators

The influence function allows to compute the asymptotic variance (Hampel et al., 1986) of dispersion estimators. The next proposition shows that for the bivariate Gaussian distributions, the asymptotic variance of covariance estimators is minimal in the independent case, and increases strictly with the absolute value of the underlying covariance. This result is similar to the one derived by Genton (1998) for scale estimators. The asymptotic variance of covariance estimators is shown to be the same, up to a multiplicative constant equal to the asymptotic variance of the underlying scale estimator in the independent case. This suggests the use of a robust estimator of scale with asymptotic variance as close as possible to the one of the MLE of scale. Consider the bivariate Gaussian distribution Φ_τ, $-1<\tau<1$, with mean zero and covariance matrix:

$$\Sigma=\begin{pmatrix}1&\tau\\ \tau&1\end{pmatrix},\tag{12}$$

and denote by Φ the standard Gaussian distribution, with density ϕ.

Proposition 3. *Let Θ be a statistical functional of covariance based on a statistical functional S of scale. The asymptotic variance of Θ at the bivariate Gaussian distribution Φ_τ is*

$$V(\Theta,\Phi_\tau)=2(1+\tau^2)V(S,\Phi),$$

where $V(S,\Phi)$ is the asymptotic variance of S at Φ.

Proof. The asymptotic variance of Θ at Φ_τ is

$$\begin{aligned}V(\Theta,\Phi_\tau)&=\iint \mathrm{IF}^2((u,v);\Theta,\Phi_\tau)\,\mathrm{d}\Phi_\tau(u,v)\\&=\frac{1}{4}\iint\left[(2+2\tau)\,\mathrm{IF}\left(\frac{u+v}{\sqrt{2+2\tau}};S,\tilde{\Phi}_+\right)-(2-2\tau)\,\mathrm{IF}\left(\frac{u-v}{\sqrt{2-2\tau}};S,\tilde{\Phi}_-\right)\right]^2\mathrm{d}\Phi_\tau(u,v).\end{aligned}$$

The change of variables $x=(u+v)/\sqrt{2+2\tau}$ and $y=(u-v)/\sqrt{2-2\tau}$ yields

$$\begin{aligned}V(\Theta,\Phi_\tau)&=\frac{1}{4}\left[(2+2\tau)^2\iint \mathrm{IF}^2(x;S,\Phi)\,\mathrm{d}\Phi(x)\mathrm{d}\Phi(y)+0\right.\\&\qquad\left.+(2-2\tau)^2\iint \mathrm{IF}^2(y;S,\Phi)\mathrm{d}\Phi(x)\,\mathrm{d}\Phi(y)\right]\\&=2(1+\tau^2)V(S,\Phi).\qquad\square\end{aligned}$$

We observe that the behavior of the asymptotic variance of correlation estimators at bivariate Gaussian distributions is opposite to the one for covariance estimators. It seems that it is maximal in the independent case, and decreases strictly with the absolute value of the underlying correlation. However, no simple proof is available, due to the much more complicated form of the influence function of correlation estimators.

We focus now on dispersion estimators based on the family of M-estimators of scale, i.e., on statistical functional $S(F)$ defined by

$$\int \chi(z/S(F))\,\mathrm{d}F(z)=0, \tag{13}$$

where χ is a real and sufficiently regular even function. The influence function of an M-estimator of scale S at F is given by (Hampel et al., 1986)

$$\mathrm{IF}(u;S,F)=\frac{\chi(u/S(F))S(F)}{B(\chi,F)} \tag{14}$$

where $B(\chi,F)=\int\chi'(z/S(F))(z/S(F))\,\mathrm{d}F(z)$. The influence function (8) of the statistical functional of covariance Θ at $\boldsymbol{F}$, based on M-estimators of scale becomes

$$\mathrm{IF}((u,v);\Theta,\boldsymbol{F})=\frac{\sigma_X\sigma_Y}{2}\left[\frac{\chi((u/\sigma_X+v/\sigma_Y)/\sigma_+)\,\sigma_+^2}{B(\chi,\tilde{F}_+)}-\frac{\chi((u/\sigma_X-v/\sigma_Y)/\sigma_-)\,\sigma_-^2}{B(\chi,\tilde{F}_-)}\right]. \tag{15}$$

Similarly, the influence function (11) of the statistical functional of correlation R at $\boldsymbol{F}$, based on M-estimators of scale becomes

$$\mathrm{IF}((u,v);R,\boldsymbol{F})=\frac{1}{\sigma_X\sigma_Y}\mathrm{IF}((u,v);\Theta,\boldsymbol{F})-\rho\left(\frac{\chi(u/\sigma_X)}{B(\chi,\tilde{F}_X)}+\frac{\chi(v/\sigma_Y)}{B(\chi,\tilde{F}_Y)}\right). \tag{16}$$

We now analyze the behavior of dispersion estimators based on some typical M-estimators of scale: MLE, MAD, and Welsch (Hampel et al., 1986). We focus on the bivariate Gaussian distribution Φ_τ, $-1<\tau<1$.

The maximum likelihood estimator (MLE) of scale at $F=\Phi$ is defined by the function $\chi(z)=z^2-1$, and yields the classical standard deviation. The covariance estimator based on the MLE corresponds to the estimator given in Eq. (2). From Eqs. (15) and (16), we get the following influence functions:

$$\mathrm{IF}((u,v);\Theta,\Phi_\tau)=uv-\tau,$$

$$\mathrm{IF}((u,v);R,\Phi_\tau)=uv-\frac{\tau}{2}(u^2+v^2).$$

Both influence functions are unbounded, which means that the dispersion estimators are not B-robust. Note that the influence function $\mathrm{IF}((u,v);R,\Phi_\tau)$ is the same as the one derived by C.L. Mallows in an unpublished manuscript cited by Devlin et al. (1975, 1976). The asymptotic variance of the dispersion estimators can be computed by integrating the square of the influence functions (Hampel et al., 1986):

$$V(\Theta,\Phi_\tau)=\int \mathrm{IF}((u,v);\Theta,\Phi_\tau)^2\,\mathrm{d}\Phi_\tau(u,v)=1+\tau^2,$$

$$V(R,\Phi_\tau)=\int \mathrm{IF}((u,v);R,\Phi_\tau)^2\,\mathrm{d}\Phi_\tau(u,v)=(1-\tau^2)^2.$$

Note that the asymptotic variance of Θ is minimal in the independent case ($\tau=0$), and must necessarily increase for the dependent data, as stated in Proposition 3. On the contrary, the asymptotic variance of R is maximal in the independent case, and decreases for dependent data.

The median absolute deviation (MAD) estimator at F is defined by the function $\chi(z)=\mathrm{sign}(|z|-q)$, where $q=F^{-1}(\frac{3}{4})$. Note that since MAD is the most B-robust estimator of scale (Hampel et al., 1986), it follows from Propositions 1 and 2 that the corresponding dispersion estimators are the most B-robust dispersion estimators among those with $\gamma_+^*(S,\tilde{F}_\pm)=\gamma_-^*(S,\tilde{F}_\pm)=\gamma^*(S,\tilde{F}_\pm)$. When $\tilde{F}_+=\tilde{F}_-$, the gross errors of the dispersion estimators are

$$\gamma^*(\Theta,\boldsymbol{F})=2\sigma_X\sigma_Y\gamma^*(MAD,\tilde{F}_+),$$

$$\gamma^*(R,\boldsymbol{F})=2(1+\rho)\gamma^*(MAD,\tilde{F}_+).$$

Note that $\gamma^*(\Theta,\boldsymbol{F})$ does not depend on the underlying covariance θ. When $\boldsymbol{F}=\Phi_\tau$, the dispersion estimators have the following influence functions:

$$\mathrm{IF}((u,v);\Theta,\Phi_\tau)=\frac{1}{8q\phi(q)}[\sigma_+^2\,\mathrm{sign}(|u+v|/\sigma_+-q)-\sigma_-^2\,\mathrm{sign}(|u-v|/\sigma_--q)],$$

$$\mathrm{IF}((u,v);R,\Phi_\tau)=\mathrm{IF}((u,v);\Theta,\Phi_\tau)-\frac{\tau}{4q\phi(q)}(\mathrm{sign}(|u|-q)+\mathrm{sign}(|v|-q)).$$

Both influence functions are bounded, which means that the dispersion estimators are B-robust. Fig. 1 depicts the influence functions $\mathrm{IF}((u,v);\Theta,\Phi_\tau)$ and $\mathrm{IF}((u,v);R,\Phi_\tau)$ for $\tau=0.3$.

Welsch's estimator of scale at $F=\Phi$ is defined by the function $\chi(z)=\frac{1}{2}-\exp(-3z^2/2)$. The corresponding dispersion estimators have the following influence functions

$$\mathrm{IF}((u,v);\Theta,\Phi_\tau)=\tfrac{4}{3}[2\tau-\sigma_+^2\exp(-\tfrac{3}{2}(u+v)^2/\sigma_+^2)+\sigma_-^2\exp(-\tfrac{3}{2}(u-v)^2/\sigma_-^2)],$$

$$\mathrm{IF}((u,v);R,\Phi_\tau)=\mathrm{IF}((u,v);\Theta,\Phi_\tau)-\frac{8\tau}{3}(1-\exp(-3u^2/2)-\exp(-3v^2/2)).$$

Both influence functions are bounded, which means that the dispersion estimators are B-robust. Fig. 2 depicts the influence functions $\mathrm{IF}((u,v);\Theta,\Phi_\tau)$ and $\mathrm{IF}((u,v);R,\Phi_\tau)$ for $\tau=0.5$.

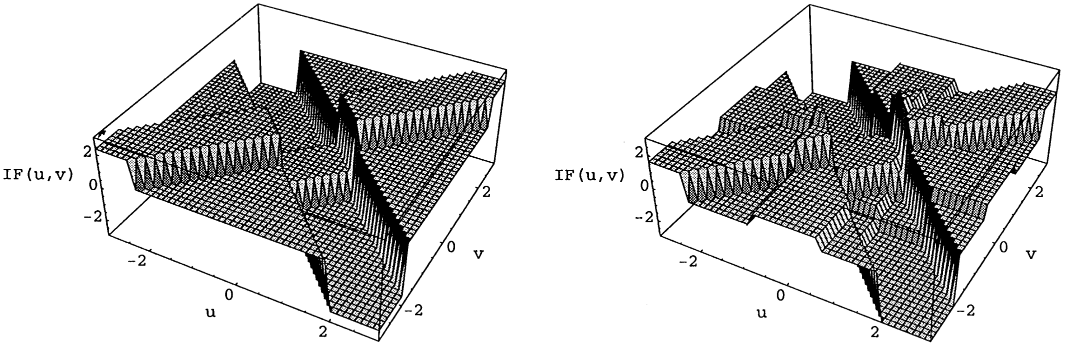

Fig. 1. The influence functions $\mathrm{IF}((u,v);\Theta,\Phi_{0.3})$ and $\mathrm{IF}((u,v);R,\Phi_{0.3})$ for the dispersion estimators based on the MAD, at the bivariate Gaussian distribution with $\tau=0.3$.

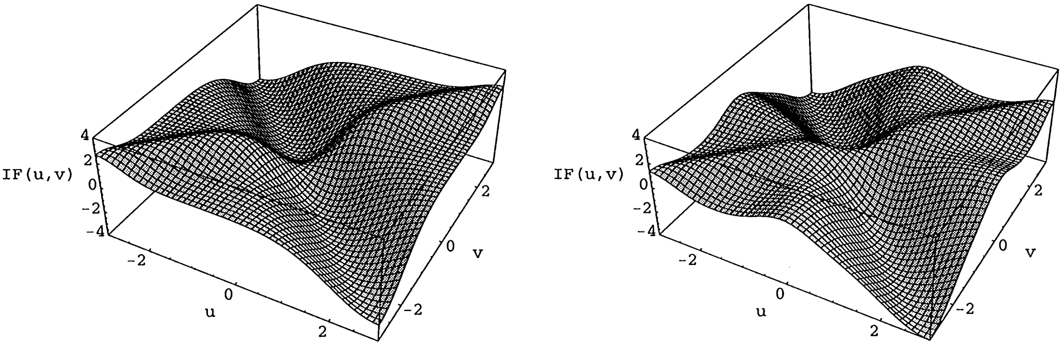

Fig. 2. The influence functions $-\mathrm{IF}((u,v);\Theta,\Phi_{0.5})$ and $\mathrm{IF}((u,v);R,\Phi_{0.5})$ for the dispersion estimators based on Welsch's scale estimator, at the bivariate Gaussian distribution with $\tau=0.5$.

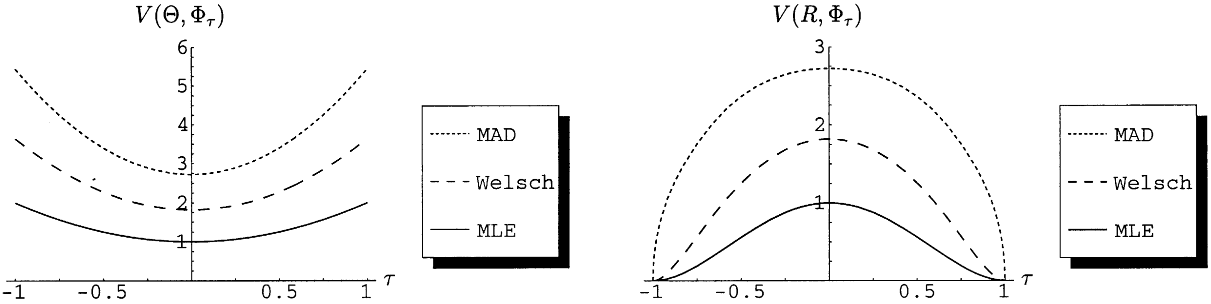

Fig. 3. The asymptotic variance of covariance estimators (top) and correlation estimators (bottom) at the bivariate Gaussian distribution Φ_τ, as a function of τ. The dispersion estimators are based on MLE, MAD, and Welsch's scale estimator. In the independent case ($\tau=0$), the variance is minimal for covariance estimators, whereas it is maximal for correlation estimators.

The asymptotic variances $V(\Theta,\Phi_\tau)$ and $V(R,\Phi_\tau)$ of covariance and correlation estimators, as a function of τ, are depicted in Fig. 3. The dispersion estimators are based on the MLE, the MAD, and Welsch's scale estimator. As shown in Proposition 3, the variance of covariance estimators is minimal in the independent case ($\tau=0$), and must necessarily increase for the dependent data. The opposite happens for correlation estimators, where the asymptotic variance is maximal in the independent case, and decreases for dependent data. As a consequence, correlation estimation is easier than covariance estimation, in the sense that it has smaller variability. However, the asymptotic variance of correlation estimators depends in a complicated way on the underlying correlation τ. For instance, consider the asymptotic variance of the correlation estimator based on Welsch's scale estimator. Straight-forward but tedious computations show that

$$V(R,\Phi_\tau)=\frac{32}{9\sqrt{7}}(4-\sqrt{7}+(8+\sqrt{7})\tau^2)+\frac{128\tau^2}{9\sqrt{16-9\tau^2}}+\frac{256\sqrt{2}\tau}{9}\left[\frac{1-\tau}{\sqrt{23+9\tau}}-\frac{1+\tau}{\sqrt{23-9\tau}}\right],$$

which is quite different from the asymptotic variance $(1-\tau^2)^2$ of the correlation estimator based on the MLE of scale. Note also that dispersion estimators based on MLE have the smallest asymptotic variance.

References

Devlin, S.J., Gnanadesikan, R., Kettenring, J.R., 1975. Robust estimation and outlier detection with correlation coefficients. Biometrika 62, 531–545.

Devlin, S.J., Gnanadesikan, R., Kettenring, J.R., 1976. Some multivariate applications of elliptical distributions. In: S. Ikeda et al. (Eds.), Essays in Probability and Statistics. Tokyo, Japan, pp. 365–393.

Fang, K., Zhang, Y., 1990. Generalized Multivariate Analysis. Springer, Berlin.

Genton, M.G., 1998. Asymptotic variance of M-estimators for dependent Gaussian random variables. Statist. Probab. Lett. 38, 255–261.

Genton, M.G., Rousseeuw, P.J., 1995. The change-of-variance function of M-estimators of scale under general contamination. J. Comput. Appl. Math. 64, 69–80.

Gnanadesikan, R., 1997. Methods for Statistical Data Analysis of Multivariate Observations, second edition. Wiley, New York.

Gnanadesikan, R., Kettenring, J.R., 1972. Robust estimates, residuals, and outlier detection with multireponse data. Biometrics 28, 81–124.

Hampel, F.R., 1974. The influence curve and its role in robust estimation. J. Amer. Statist. Assoc. 69, 383–393.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. Robust Statistics: The Approach based on Influence Functions. Wiley, New York.

Huber, P.J., 1981. Robust Statistics. Wiley, New York.

Kano, Y., 1994. Consistency property of elliptical probability density functions. J. Multivariate Anal. 51, 139–147.

Ma, Y., Genton, M.G., 1998. Highly robust estimation of the autocovariance function. J. Time Ser. Anal., submitted for publication.

Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. Multivariate Analysis. Academic Press, New York.