

Finding Extreme Values and Extreme Points of a Multivariate Function *

C. C. Y. Dorea

Abstract: Let f be a k -variate function defined on $\Omega \subset R^d$ and consider the problem of estimating the extreme values of f and the corresponding extreme points in Ω . Conditions that will assure common extreme points for the coordinate functions $\{f_j\}_{j=1,\dots,k}$ will be discussed. Also test for the asymptotic independence under weak convergence of the coordinate functions will be presented.

Key words: Multivariate extremes, regular variation, limiting distribution, asymptotic independence.

1. Introduction.

Since their introduction by Fisher and Tippett (1928), univariate extreme value distributions have been extensively studied, perhaps most notably by Gnedenko (1943). Results for the multivariate case, obtained by a number of authors, have been summarized by Galambos (1978). A representation of bivariate maximal extreme value distribution H that asymmetrically involve the marginal distributions was obtained by Sibuya (1960), and Berman (1962) obtains necessary and sufficient conditions for a bivariate distribution F to be in the domain of attraction of such an H . Sibuya also introduces the notion of "dependence function" which was successfully used by Tiago de Oliveira (1963-1975) to obtain the structure of bivariate extreme distributions. For another approach to this problem see de Haan and Resnick (1977), they make use of the theory of max infinite divisible distributions and the notion of regularly varying functions. An extensive treatment of this approach can be found in Resnick (1987). An approach that avoids the use of the dependence function to characterize the domains of attraction was derived by Marshall and Olkin (1983).

We consider the problem of estimating the extreme values and the extreme points of a multivariate function. Our approach makes use of the notion of regularly varying functions of de Hann and Resnick and conditions derived from Marshall and Olkin.

Let $f = (f_1, \dots, f_k)$ be a k -variate function defined on some measurable domain Ω of R^d and consider the problem of estimating " $\min_{x \in \Omega} \{f_1(x), \dots, f_k(x)\}$ " (or $\max_{x \in \Omega} \{f_1(x), \dots, f_k(x)\}$). Clearly, ordering multivariate data is ambiguous and determining extreme values is subjective. This problem can be partially solved

* Talk presented at V Congresso Latino-Americano de Probabilidade e Estatística Matemática, USP, São Paulo, June 1993.

by analysing the asymptotic independence of the coordinate functions under the weak convergence. We will see that if the coordinate function are asymptotically independent then the extreme values have to be estimated separately. On the other hand if the coordinate functions are asymptotically dependent then they necessarily have common extreme points and the extreme value can be estimated jointly.

Let's consider the following related problem: for ξ_1, ξ_2, \dots independent and identically distributed (i.i.d.) random vectors with $P(\Omega) > 0$, analyse the asymptotic behavior of the random vector $\left(\frac{\min\{f_\ell(\xi_1), \dots, f_\ell(\xi_n)\} - a_{\ell,j}}{b_{\ell,j}} \right)_{\ell=1, \dots, k}$ as $n \rightarrow \infty$. Where $a_{\ell,j} > 0$ and $b_{\ell,j}$ are stabilizing constants with $b_{\ell,j} > 0$. If it converges weakly to a nondegenerate distribution then the interaction of the minimum points of the coordinate functions can be explicitly displayed. The minimum points of the coordinate functions are defined by:

$$M(f_j) = \{x : x \in \Omega, f_j(x) = \min_{x \in \Omega} (f_j(x))\}, \quad j = 1, \dots, k.$$

For completeness in section 2 we treat the univariate case by detailing the estimation of the minimum points. For a more extensive treatment we refer the reader to Dorea (1987).

In section 3 we present the bivariate case and indicate the treatment for the general multivariate case. As a general reference for this section see Dorea (1993) and Dorea and Mizaki (1993).

2. The univariate case.

Let ξ_1, ξ_2, \dots be i.i.d. random variables with a common distribution G such that $P(\Omega) > 0$. To estimate the global minimum of f on Ω :

$$y = \min\{f(x) : x \in \Omega\} \tag{1}$$

we analyse the asymptotic behavior of $Y(n) = \min\{f(\xi_1), \dots, f(\xi_n)\}$. Let F be the common distribution of the i.i.d. random variables $f(\xi_1), f(\xi_2), \dots$. From the classical results of Fisher-Tippett and Gnedenko we have the necessary and sufficient conditions for the existence of norming constants $b_n > 0$ and a_n such that for all continuity points x of H we have

$$\lim_{n \rightarrow \infty} P(Y(n) \leq b_n x + a_n) = H(x) \tag{2}$$

where H is a nondegenerate distribution. That is F is in the domain of attraction of H . In short, we write $F \in \mathcal{D}(H)$ or $f \in \mathcal{D}(H)$. Moreover, H is

necessarily one of the three classes: for a constant $\alpha > 0$,

$$\psi_\alpha(x) = \begin{cases} 1 - \exp(-(-x)^{-\alpha}) & , x < 0, \\ 1 & , x \geq 0, \end{cases}$$

$$\phi_\alpha(x) = \begin{cases} 0, & x < 0 \\ 1 - \exp(-x^\alpha), & x \geq 0 \end{cases}$$

and

$$\Lambda(x) = 1 - \exp(-e^x).$$

It is easy to verify that the Cauchy distribution is in $\mathcal{D}(\psi_\alpha)$ while the uniform and normal distribution are in $\mathcal{D}(\phi_\alpha)$ and $\mathcal{D}(\Lambda)$ respectively. We will consider the specific case of ϕ_α , in which we have \mathbf{y} finite. And we may take

$$a_n = \mathbf{y} \text{ and } b_n = \sup\{x : F(x) \leq \frac{1}{n}\} - \mathbf{y}. \quad (3)$$

Remark 1. Treating only the case ϕ_α is not restrictive since by monotonic transformations one can reduce ψ_α and Λ cases to the ϕ_α case. For that let Y such that $Y \in \mathcal{D}(\psi_\alpha)$ then $-\frac{1}{Y} \in \mathcal{D}(\phi_\alpha)$. Also, if $Y \in \mathcal{D}(\Lambda)$ and the corresponding norming constants $b_n \rightarrow 0$ then $e^Y e^{\mathbf{y}} \in \mathcal{D}(\phi_\alpha)$. Equivalently, regarding the original function f one would be estimating the minimum of $-\frac{1}{f}$ or $e^f + e^{\mathbf{y}}$.

Our problem now reduces to finding conditions on f that will assure $f(\xi_1) \in \mathcal{D}(\phi_\alpha)$. Without loss of generality we will assume that $\Omega \subset \mathbb{R}$, $\Omega = [0, 1]$ and that ξ_1 is uniformly distributed on $[0, 1]$. Note that if ξ_1 has a continuous distribution G on Ω and

$$G^-(\mathbf{y}) = \sup\{t : G(t) \leq \mathbf{y}\}, \quad 0 < \mathbf{y} < 1. \quad (4)$$

Then for U uniformly distributed on $[0, 1]$ we have $G^-(U)$ with distribution G . And it is enough to study the function $\varphi = f(G^-)$.

Now assume that f is defined on the unit interval I and that the minimum $\mathbf{y} = \min_{x \in I} \{f(x)\}$ is finite. Let M denote the set of the minimum points of f :

$$M = \{x : x \in I, f(x) = \mathbf{y}\}. \quad (5)$$

We say that f satisfies Condition 1 if: for some $\delta > 0$ there exists a δ -varying function $v(t)$ such that for each $x_0 \in M$ and all $x \neq 0$ the following limit exists (possibly ∞):

$$R(x_0, x) = \lim_{t \rightarrow 0} \frac{f(x_0 + tx) - \mathbf{y}}{v(t)} \quad (6)$$

Note that if the minimum point x_0 is one of the endpoints 0 or 1 we interpret (6) as holding for all $x > 0$ or $x < 0$ respectively. We say that a function $v(t)$, $t > 0$ is δ -varying if for all $x > 0$ we have $\lim_{t \downarrow 0} (v(tx)/v(t)) = x^\delta$ (see also de Haan (1971) and Resnick (1987)).

Relative to the norming function v we now associate to each minimum point x_0 a measure of its minimality contact by defining a measure μ on M . For $x_0 \in M$ let

$$\mu(x_0) = (R(x_0, -1))^{-1/\delta} + (R(x_0, 1))^{-1/\delta}, \quad (7)$$

where $\mu(x_0) = \infty$ if either $R(x_0, 1)$ or $R(x_0, -1) = 0$ and $\mu(x_0) = 0$ if $R(x_0, 1) = R(x_0, -1) = \infty$. For $E \subset M$ define $\mu(E) = \infty$ if for some $x_0 \in E$, $\mu(x_0) = \infty$; otherwise let $\mu(E) = \sum_{x \in E, 0 < \mu(x) < \infty} \mu(x)$. Theorem 1 below shows that $\mu(M)$ characterizes the asymptotic behavior of

$$Y(n) = \min\{f(U_1), \dots, f(U_n)\}$$

where U_1, U_2, \dots are independent and uniformly distributed on I . For its proof as well as the proof of the theorem 2 see Dorea (1987).

Theorem 1. Under Condition 1 we have for all $x > 0$

$$\lim_{n \rightarrow \infty} P \left(\frac{Y(n) - \mathbf{y}}{v(1/n)} \leq x \right) = 1 - \exp\{-\mu(M)x^\alpha\}, \quad (8)$$

where $\alpha = 1/\delta$ and (8) should be interpreted as $\frac{Y(n) - \mathbf{y}}{v(1/n)}$ diverging to ∞ in probability if $\mu(M) = 0$ and converging to the degenerate distribution if $\mu(M) = \infty$.

Theorem 2 below can be viewed as a characterization of M relative to the weak convergence.

Theorem 2. If Condition 1 is satisfied and $\mu(M) < \infty$ then there are at most finitely many minimum points x_0 and also $\mu(x_0) < \infty$ for each of them.

Our next result, theorem 3, illustrates the role played by the measure of minimality contact μ . It states that given that $Y(n)$ falls within an ϵ -neighborhood

of \mathbf{y} the asymptotic conditional probability of U_j (for some $1 \leq j \leq n$) to fall within a neighborhood of the minimum point x_0 is $\mu(x_0)/\mu(M)$.

Theorem 3. Let Condition 1 be verified and assume that $0 < \mu(M) < \infty$. Then there exists a norming function $u(\cdot)$ such that for each $x_0 \in M$ there will be constants k^- and k^+ with

$$\lim_{\epsilon \downarrow 0} \lim_{n \rightarrow \infty} P \left(H_n(\epsilon) \mid Y(n) \leq v\left(\frac{1}{n}\right)\epsilon + \mathbf{y} \right) = \frac{\mu(x_0)}{\mu(M)}, \quad (9)$$

$$\text{where } H_n(\epsilon) = \bigcup_{1 \leq j \leq n} \left\{ -k^- \epsilon^{1/\delta} \leq \frac{(x_j - x_0)}{u(1/n)} \leq k^+ \epsilon^{1/\delta} \right\}.$$

Remark 2. (a) Note that each minimum point can be estimated by theorem 3 provided one can estimate δ and $\mu(M)$. And this can be done by considering the order statistics $Z_{(1)}, Z_{(2)}, \dots$ of the i.i.d. random variables $\{Z_k\}_{k \geq 1}$ with $Z_k = f(U_k)$. Let $k(n) \rightarrow \infty$ with $\frac{k(n)}{n} \rightarrow 0$ we have

$$L_n = \log \frac{Z_{(k(n))} - Z_{(3)}}{Z_{(3)} - Z_{(2)}} / \log k(n) \xrightarrow{P} \delta. \quad (10)$$

Furthermore, we have

$$\frac{1}{k(n)} \left(\frac{Z_{(k(n))} - Z_{(1)}}{v(1/n)} \right)^{1/L_n} \xrightarrow{P} \frac{1}{\mu(M)}. \quad (11)$$

(b) Assume now the case where ξ_1, ξ_2, \dots are i.i.d. with a common absolutely continuous distribution G and $P(\Omega) = 1$ with $\Omega \subset R$. Let G^- be defined by (3). We say that f and G satisfy Condition 2 if: $G'(x) > 0$ a.e. on Ω and for some $\delta > 0$ there exists a δ -varying function $v(t)$ such that for each $x_0 \in M$ and all $x \neq 0$ the following limit exists

$$R_G(x_0, x) = \lim_{t \downarrow 0} \frac{f(G^-(G(x_0) + tx)) - \mathbf{y}}{v(t)}. \quad (12)$$

Similarly to (7) one defines measure μ_G on M . For $x_0 \in M$

$$\mu_G(x_0) = (R_G(x_0, -1))^{-1/\delta} + (R_G(x_0, 1))^{-1/\delta}. \quad (13)$$

Under Condition 2 we have (8) with $\alpha = 1/\delta$.

(c) For the multidimensional case $\Omega \subset R^d$. If $\Omega = I^d$ the unit cube (6) becomes

$$R(x_0, x) = \lim_{t \downarrow 0} \frac{f(x_0 + tx) - \mathbf{y}}{v(t)}$$

holding for all $x \in R^d$, $x \neq 0$. As for the measure of minimality contact one can show that in the one dimensional case we have $\mu(x_0) = m\{x : R(x_0, x) \leq 1\}$ being m the Lebesgue measure. For the d -dimensional case take $\mu(x_0) = m_d\{x : R(x_0, x) \leq 1\}$ being m_d the Lebesgue measure on R^d . And theorem 1 follows with $v((\frac{1}{n})^{1/d})$ in place of $v(\frac{1}{n})$. For general $\Omega \subset R^d$ and assuming the random vectors ξ_1, ξ_2, \dots have independent components, that is, ξ_1, ξ_2, \dots are i.i.d. with common distribution $G = G_1 G_2 \dots G_d$ the treatment is analogous to remark 2(b). In this case for $z = (z_1, \dots, z_d)$ take

$$G^-(z) = (\inf\{t : G_i(t) \leq z_i\}, \quad i = 1, \dots, d).$$

3. The multivariate case

To avoid heavy notation we will present the bivariate case. The general case can be similarly handled. Let $(f_1, f_2) = (f, g) : \Omega \rightarrow R^2$. Let

$$\mathbf{y} = \min_{x \in \Omega} \{f(x)\} \quad \text{and} \quad \mathbf{z} = \min_{x \in \Omega} \{g(x)\}. \quad (14)$$

Just as in the univariate case one can assume that $\Omega = I$ the unit interval and that both \mathbf{y} and \mathbf{z} are finite. We say that a distribution F is of type ϕ_α if for some constant $a > 0$ it is of the form:

$$F(x) = \begin{cases} 0 & , x < 0 \\ 1 - \exp(-ax^\alpha) & , x \geq 0. \end{cases}$$

Assume that H is a bivariate distribution with marginals F and G of type ϕ_α and ϕ_β respectively. We say that (f, g) is in the domain of attraction of H , $(f, g) \in \mathcal{D}(H)$, if for U_1, U_2, \dots i.i.d. uniformly distributed on I , there exists constants $a_n, b_n > 0, c_n$ and $d_n > 0$ such that for $Y(n) = \min\{f(U_1), \dots, f(U_n)\}$ and $Z(n) = \min\{g(U_1), \dots, g(U_n)\}$ we have

$$\lim_{n \rightarrow \infty} P(Y(n) \leq b_n x + a_n, Z(n) \leq d_n y + c_n) = H(x, y) \quad (15)$$

for all continuity points (x, y) of H . Clearly if (15) holds we have $f \in \mathcal{D}(F)$ and $g \in \mathcal{D}(G)$. Since we are taking F and G of type ϕ_α and ϕ_β one can take $a_n = \mathbf{y}$, $c_n = \mathbf{z}$,

$$b_n = \sup\{x : R(x) \leq \frac{1}{n}\} - \mathbf{y}$$

and

$$d_n = \sup\{x : S(x) \leq \frac{1}{n}\} - \mathbf{z}.$$

Where R and S are the marginal distributions of $(f(U_1), g(U_1))$. Let L be the joint distribution of $(f(U_1), g(U_1))$ and $\bar{L}(x, y) = P(f(U_1) > x, g(U_1) > y)$ then a sufficient condition for (15) to hold is that there exists a δ -varying function $v(\cdot)$ with $\delta = 1/\alpha$ such that the following limit exists,

$$\lim_{t \downarrow 0} \frac{1 - \bar{L}(\mathbf{y} + t\mathbf{x}, \mathbf{z} + s(t)\mathbf{y})}{v(t)} \tag{16}$$

for all (x, y) such that $\bar{H}(x, y) > 0$. Where $s(t) = S^-(R(\mathbf{y} + t)) - \mathbf{z}$ with $S^-(u) = \sup\{t : S(t) \leq u\}$, $0 < u < 1$. Clearly if (16) is satisfied then f and g satisfy condition 1 with $\delta = 1/\alpha$ and $\gamma = 1/\beta$ respectively.

Now let $M(f)$ and $M(g)$ be the set of minimum points of f and g respectively, that is,

$$M(f) = \{x : x \in I, f(x) = \mathbf{y}\} \text{ and } M(g) = \{x : x \in I, g(x) = \mathbf{z}\}. \tag{17}$$

Theorem 4. Under condition (16) we have $H = FG$ if and only if $M(f) \cap M(g) = \phi$.

Note that if $M(f) \cap M(g) = \phi$ then $Y(n)$ and $Z(n)$ are asymptotically independent and \mathbf{y} and \mathbf{z} have to be estimated separately. In case $M(f) \cap M(g) \neq \phi$ one can construct confidence region for (\mathbf{y}, \mathbf{z}) based on the limiting distribution H and on the set of common minimum points. Therefore an explicit representations of H would be helpful. Theorem 5 below shows that such representation is possible. The following examples motivate our results.

Examples. Let μ and w be respectively the measure of minimality contact of f and g as defined in section 2. (a) If $f(x) = |1 - 2x|$ and $g(x) = (x - \frac{1}{2})^2$ then $M(f) = M(g) = \{\frac{1}{2}\}$, $\mu(1/2) = 1$ and $\nu(\frac{1}{2}) = 2$. We have $f \in \mathcal{D}(1 - e^{-x})$, $g \in \mathcal{D}(1 - e^{-2\sqrt{y}})$ and $(f, g) \in \mathcal{D}((1 - e^{-x}) \wedge (1 - e^{-2\sqrt{y}}))$ where $x > 0$ and $y > 0$. (b) If $f(x) = |1 - 2x|$ and $g(x) = x^2$ then $M(f) = \{\frac{1}{2}\}$, $M(g) = \{0\}$, $M(f) \cap M(g) = \phi$, $\mu(\frac{1}{2}) = 1$ and $\nu(0) = 2$. We have $f \in \mathcal{D}(1 - e^{-x})$, $g \in \mathcal{D}(1 - e^{-2\sqrt{y}})$ and $(f, g) \in \mathcal{D}((1 - e^{-x})(1 - e^{-2\sqrt{y}}))$. (c) If $f(x) = x$, $0 \leq x \leq \frac{1}{4}$; $f(x) = |\frac{1}{2} - x|$, $\frac{1}{4} < x \leq 1$ and $g(x) = x^2$, $0 \leq x \leq \frac{1}{4}$; $g(x) = (\frac{1}{x} - x)^2$, $\frac{1}{4} < x \leq 1$. Then $M(f) = \{0, \frac{1}{2}\}$, $M(g) = \{0, \frac{1}{3}\}$, $M(f) \cap M(g) = \{0\}$, $\mu(0) = 1$, $\mu(\frac{1}{2}) = 2$, $\nu(0) = 1$ and $\nu(\frac{1}{3}) = 2$. We have $f \in \mathcal{D}(1 - e^{-3x})$ and $g \in \mathcal{D}(1 - e^{-3\sqrt{y}})$. Note that the common minimum point 0 indicates the dependence of f and g and its contribution towards the limiting distributions F and G can be expressed as $F_D(x) = 1 - e^{-x}$ and $G_D(x) = 1 - e^{-\sqrt{y}}(\mu(0) = \nu(0) = 1)$. As for the

independent minimum point $\{\frac{1}{2}\}$ of f with $u(\frac{1}{2}) = 2$ we have $F_I = 1 - e^{-2x}$. And for the independent minimum point $\{\frac{1}{3}\}$ of G with $\nu(\frac{1}{3}) = 2$ we have $1 - e^{-2\sqrt{y}}$. And we have $(f, g) \in \mathcal{D}(H)$ where H can be expressed as: if $F_D \leq G_D$ then $H = F_I G + F - F_I$; if $G_D \leq F_D$ then $H = G_I F + G - G_I$.

Theorem 5. Assume that f and g satisfy Condition 1 for some $\delta > 0$ and $\gamma > 0$ and that $(f, g) \in \mathcal{D}(H)$. If F and G are the marginal distributions of H then H has the following representation:

$$\begin{aligned} H &= F_I G + (F - F_I) \quad \text{if } F_D \leq G_D \\ &= G_I F + (G - G_I) \quad \text{if } G_D \leq F_D \end{aligned} \quad (18)$$

where for $x > 0$ and $y > 0$, $F_I(x) = 1 - \exp(-\mu(I_f)x^\alpha)$, $F_D(x) = 1 - \exp(-\mu(D)x^\alpha)$, $G_I(y) = 1 - \exp(-\nu(I_g)y^\beta)$, $G_D(y) = 1 - \exp(-\nu(D)y^\beta)$, $\alpha = \frac{1}{\delta}$, $\beta = \frac{1}{\gamma}$, $D = M(f) \cap M(g)$, $I_f = M(f) \setminus D$, $I_g = M(g) \setminus D$, μ and ν are respectively the measure of minimality contact of f and g as defined in (7).

Remark 3. (1) If $D = \phi$, that is f and g have no common minimum points then $F_I = F$ and $G_I = G$ so that $H = FG$.

(2) If $M(f) \equiv M(g)$ then $F_D = F$ and $G_D = G$ so that $H = F \wedge G$.

(3) For the general case $f = (f_1, \dots, f_k) : \Omega \rightarrow R^k$, (18) becomes for more complex involving all possible situations. But the treatment is similar.

Corollary. Under conditions of theorem 5 if for some (x_0, y_0) such that $0 < \overline{F(x_0)} < 1$ and $0 < G(x_0) < 1$ we have $H(x_0, y_0) = F(x_0)G(x_0)$ then $H(x, y) = F(x)G(y)$ for all $(x, y) \in R^2$.

This suggests the following test for asymptotic independence (see Dorea and Mizaki (1993)): under the assumptions of theorem 5, if for some (x_0, y_0) with $x_0 > 0$ and $y_0 > 0$ we have:

$$\frac{n\tau_n(f, g)}{\tau_n(f)\tau_n(g)} \leq 1 \quad (19)$$

then $Y(n)$ and $Z(n)$ are asymptotically independent. Where $\tau_n(f) = \sum_{j=1}^n \eta_j$

with $\eta_j = 1$ if $f(U_j) \leq b_n x_0 + a_n$ and $\eta_j = 0$ otherwise; $\tau_n(g) = \sum_{\ell=1}^n \rho_\ell$ with $\rho_\ell = 1$ if $g(U_\ell) \leq d_n y_0 + c_n$ and $\rho_\ell = 0$ otherwise; and $\tau_n(f, g) = \sum_{j=1}^n \theta_j$ with $\theta_j = \eta_j \rho_j$.

References:

- 1 Berman, S.M. (1962) - "Convergence to bivariate limiting extreme value distributions", Annals of Institute Statistical Mathematics, vol. 13, p. 217-223.
- 2 Dorea, C.C.Y. (1987) - "Estimation of the extreme values and the extreme points", Annals of Institute Statistical Mathematics, vol. 39, No. 1, p. 37-48.
- 3 Dorea, C.C.Y. (1993) - "Estimation of the extreme points of a multivariate functions", to appear.
- 4 Dorea, C.C.Y. and Mizaki, E.S. (1993) - "Asymptotic independence tests for bivariate extremes", Acta Mathematica Hungarica, to appear.
- 5 Galambos, J. (1978) - "The asymptotic theory of extreme order statistics", John Wiley and Sons, New York.
- 6 de Haan, L. (1971) - "A form of regular variation and its application to the domain of attraction of the double exponential", Z. Wahrscheinlichkeitstheorie verw. Gebiete, vol. 17, p. 241-252.
- 7 de Haan, L. and Resnick, S.I. (1977) - "Limit theory for multivariate sample extremes", Z. Wahrscheinlichkeitstheorie verw. Gebiete, vol. 40, p. 317-337.
- 8 Marshall, A.W. and Olkin, I. (1983) - "Domains of attraction of multivariate extreme value distributions", The Annals of Probability, vol. 11, No. 1, p. 168-177.
- 9 Resnick, S.I. (1987) - "Extreme values, regular variation and point processes", Springer Verlag, New York.

Chang C.Y. Dorea

Departamento de Matemática

Universidade de Brasília

70910 Brasília-DF

Brasil