Res. Lett. Inf. Math. Sci., 2005, Vol.**8**, pp97-103
Available online at http://iims.massey.ac.nz/research/letters/

Continuous canonical correlation analysis

C.M. Cuadras

Departament d'Estadística
Universitat de Barcelona
Diagonal, 645
08028 Barcelona (Spain)

Given a bivariate distribution, the set of canonical correlations and functions is in general finite or countable. By using an inner product between two functions via an extension of the covariance, we find all the canonical correlations and functions for the so-called Cuadras-Augé copula and prove the continuous dimensionality of this distribution.

1 Introduction

Let H be a bivariate cdf with marginals F, G. Suppose that the measure $dH(x,y)$ is absolutely continuous with respect to $dF(x)dG(y)$ and that Pearson's contingency coefficient ϕ^2 defined by

$$\phi^2 + 1 = \int_a^b \int_c^d (dH(x,y))^2/(dF(x)dG(y))$$

is finite. Then the following expansion holds

$$dH(x,y) - dF(x)dG(y) = \sum_{n\geq 1} \rho_n a_n(x) b_n(y) dF(x)dG(y), \tag{1}$$

where ρ_n are canonical correlations, ordered in descending order, and $a_n(x), b_n(y)$ are canonical variables (Lancaster, 1969). Thus $\rho_1 = \rho(a_1(X), b_1(Y))$ is the maximum correlation between a function of X and a function of Y, $\rho_2 = \rho(a_2(X), b_2(Y))$ is the maximum correlation constrained to the functions with zero correlation with $a_1(X)$ and $b_1(Y)$, etc.

Cuadras (2002a) proved the formula giving the covariance between two functions of bounded variation in terms of the cdf's. If $H(x,y)$ is symmetric and positive-quadrant dependent, i.e., $H(x,y) \geq F(x)G(y)$, and $F \equiv G$, this formula gives rise to define the following inner product $\langle \alpha, \beta \rangle_H$ between two functions α, β with common rank $[a,b]$:

$$\begin{aligned} \langle \alpha, \beta \rangle_H &= Cov(\alpha(X), \beta(Y)) \\ &= \int_a^b \int_a^b (H(x,y) - F(x)F(y))\, d\alpha(x) d\beta(y). \end{aligned} \tag{2}$$

Cuadras (2002b) expressed (1) in terms of cdf's

$$H(x,y) - F(x)G(y) = \sum_{n \geq 1} \rho_n \int_a^b L(x,s) da_n(s) \int_c^d M(t,y) db_n(t), \tag{3}$$

where $L(x,s) = \min\{F(x), F(s)\} - F(s)F(t)$, $M(t,y) = \min\{G(t), G(y)\} - G(t)G(y)$. Making an analogy with correspondence analysis, the number of distinct canonical correlations determine the dimensionality of H, i.e., the dimension of H is $\#(\rho_n)$, see Cuadras et al. (1999). In general, this dimension is finite or countable, but can be continuous.

Cuadras and Augé (1981) proposed the bivariate distribution

$$H_\theta(x,y) = \min\{F(x), G(y)\}^\theta (F(x)G(y))^{1-\theta}, \quad 0 \leq \theta \leq 1,$$

where the marginals are F, G. The uniform transformation $U = F(X), V = G(Y)$, provides the so-called Cuadras-Augé copula

$$C_\theta(u,v) = \min\{u,v\}^\theta (uv)^{1-\theta}, \quad 0 \leq \theta \leq 1,$$

which is the survival copula of the Marshall-Olkin distribution (Nelsen, 1999). See Ruiz-Rivas and Cuadras (1988) and Genest and Plante (2003) for further aspects.

Suppose that the cdf of the random vector (U,V) is C_θ. If $\mathcal{H}_1$ is the Heaviside distribution

$$\mathcal{H}_1(x) = 0 \quad \text{if } x < 1, \; \mathcal{H}_1(x) = 1 \quad \text{if } x \geq 1,$$

it can be proved (Cuadras, 2002a) that the first canonical correlation or maximum correlation between a function of U and a function of V is

$$\rho(\mathcal{H}_1(U), \mathcal{H}_1(V)) = \max_\varphi \rho(\varphi(U), \varphi(V)) = \theta.$$

We generalize this result by extending the canonical correlation analysis to the continuous case and proving the continuous dimensionality for this copula. This approach is comparable to the continuous extension of multidimensional scaling (Cuadras and Fortiana, 1995).

2 Eigenanalysis

Let us find the eigenvalues and eigenfuctions for the covariance kernels

$$K_\theta(u,v)=\min\{u,v\}^\theta(uv)^{1-\theta}-uv \quad \text{and} \quad L(u,v)=\min\{u,v\}-uv,$$

related to the Cuadras-Augé copula. Note that $L=K_1$.

A function ϕ is an eigenfunction of K_θ with respect to L with eigenvalue λ if

$$\int_0^1 K_\theta(u,v)d\phi(v)=\lambda\int_0^1 L(u,v)d\phi(v).$$

Let us define

$$\mathcal{H}_{\gamma,\varepsilon}(x)=\mathcal{H}_{\gamma^-}(x)-\frac{\gamma}{\gamma+\varepsilon}\mathcal{H}_{(\gamma+\varepsilon)^+}(x),$$

where

$$\begin{aligned}\mathcal{H}_{\gamma^-}(x)&=0 \quad \text{if } x<\gamma,\ \mathcal{H}_{\gamma^-}(x)=1 \quad \text{if } x\geq\gamma,\\ \mathcal{H}_{\gamma^+}(x)&=0 \quad \text{if } x\leq\gamma,\ \mathcal{H}_{\gamma^+}(x)=1 \quad \text{if } x>\gamma.\end{aligned}$$

Theorem 1. *The set $(\phi_\gamma,\lambda_\gamma)$ of eigenfunctions and eigenvalues of K_θ with respect to L is given by*

$$\phi_\gamma=\lim_{\varepsilon\to 0}\mathcal{H}_{\gamma,\varepsilon},\qquad \lambda_\gamma=\theta\gamma^{1-\theta},\quad 0\leq\gamma\leq 1,$$

where ϕ_γ is the indicator of γ, i.e., $\phi_\gamma(x)=0$ if $x\neq\gamma$, and $\phi_\gamma(\gamma)=1$.

Proof. We have

$$d\mathcal{H}_{\gamma,\varepsilon}(x)=d\mathcal{H}_{\gamma^-}(x)-\frac{\gamma}{\gamma+\varepsilon}d\mathcal{H}_{(\gamma+\varepsilon)^+}(x),$$

with $d\mathcal{H}_{\gamma^-}(x)=d\mathcal{H}_{\gamma^+}(x)=0$ for $x\neq\gamma$, $d\mathcal{H}_{\gamma^-}(\gamma)=1$ and

$$\int_0^1 v^\alpha d\mathcal{H}_{\gamma^-}(v)=\int_0^1 v^\alpha d\mathcal{H}_{\gamma^+}(v)=\gamma^\alpha.$$

For $0<u<\gamma$,

$$\begin{aligned}\int_0^1 K_\theta(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v)&=\int_0^u K_\theta(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v)+\int_u^1 K_\theta(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v)\\&=0+u\int_0^1(v^{1-\theta}-v)d\mathcal{H}(\gamma,\varepsilon)\\&=u\{\gamma^{1-\theta}-\gamma-\tfrac{\gamma}{\gamma+\varepsilon}[(\gamma+\varepsilon)^{1-\theta}-(\gamma+\varepsilon)]\}\\&=u[\gamma^{1-\theta}-\gamma(\gamma+\varepsilon)^{-\theta}].\end{aligned}$$

For $\gamma\leq u\leq\gamma+\varepsilon$, where $\varepsilon>0$ is arbitrarily small,

$$\begin{aligned}\int_0^1 K_\theta(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v)&=\int_\gamma^u(vu^{1-\theta}-uv)d\mathcal{H}_{\gamma,\varepsilon}(v)+\int_u^{\gamma+\varepsilon}u(v^{1-\theta}-v)d\mathcal{H}_{\gamma,\varepsilon}(v)\\&=(u^{1-\theta}-u)\gamma+u(\gamma-\gamma(\gamma+\varepsilon)^{-\theta})\\&=\gamma[u^{1-\theta}-u(\gamma+\varepsilon)^{-\theta}].\end{aligned}$$

For $0<\gamma+\varepsilon<u$,

$$\begin{aligned}\int_0^1 K_\theta(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v) &= \gamma^\theta(u\gamma)^{1-\theta}-u\gamma-\tfrac{\gamma}{\gamma+\varepsilon}[(\gamma+\varepsilon)^\theta u^{1-\theta}-u(\gamma+\varepsilon)]\\ &= \gamma(u^{1-\theta}-u)-\gamma(u^{1-\theta}-u)\\ &= 0.\end{aligned}$$

In particular, if $\theta=1$, $K_\theta=L$ and

$$\int_0^1 L(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v)=\begin{cases} u[1-\gamma(\gamma+\varepsilon)^{-1}], & \text{if } 0<u<\gamma,\\ \gamma[1-u(\gamma+\varepsilon)^{-1}], & \text{if } \gamma\le u\le\gamma+\varepsilon,\\ 0, & \text{if } \gamma+\varepsilon<u<1.\end{cases}$$

Thus we have

$$\int_0^1 K_\theta(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v)=\lambda_\gamma(\varepsilon)\int_0^1 L(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v), \tag{4}$$

uniformly in $u\notin[\gamma,\gamma+\varepsilon]$, for

$$\begin{aligned}\lambda_\gamma(\varepsilon) &= \frac{\gamma^{1-\theta}-\gamma(\gamma+\varepsilon)^{-\theta}}{1-\gamma/(\gamma+\varepsilon)}\\ &= \frac{(\gamma+\varepsilon)\gamma^{1-\theta}-\gamma(\gamma+\varepsilon)^{1-\theta}}{\varepsilon}.\end{aligned}$$

Finally the interval $[\gamma,\gamma+\varepsilon]$ degenerates to γ as $\varepsilon\to 0$ and $\lim_{\varepsilon\to 0}\lambda_\gamma(\varepsilon)=\theta\gamma^{1-\theta}$. □

Remark 1. *Direct integration using the distribution $\mathcal{H}_\gamma(x)=0$ if $x<\gamma$, $\mathcal{H}_\gamma(x)=1$ if $x\ge\gamma$, cannot give the eigenequation (4) for some λ_γ, except for $\gamma=1$. This trouble can be overcome by using the limit*

$$\begin{aligned}\lim_{\varepsilon\to 0}\lambda_\gamma(\varepsilon) &= \lim_{\varepsilon\to 0}\frac{\int_0^1 K_\theta(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v)}{\int_0^1 L(u,v)d\mathcal{H}_{\gamma,\varepsilon}(v)}\\ &= \theta\gamma^{1-\theta}.\end{aligned}$$

if $0<u\le\gamma$ and setting $\frac{0}{0}=\theta\gamma^{1-\theta}$ if $\gamma<u<1$.

3 Canonical analysis

Let us use the notation $\langle\phi,\psi\rangle_\theta=Cov(\phi(U),\psi(V))$, $\langle\phi,\phi\rangle_1=||\phi||_1^2=Var(\phi(U))$, where the inner product $\langle\cdot,\cdot\rangle_\theta\equiv\langle\cdot,\cdot\rangle_{C_\theta}$ has been defined in (2). Thus, given two real functions $\phi,\psi\in\ell^2([0,1])$, the squared correlation between $\phi(U),\psi(V)$ can be written as

$$\begin{aligned}\rho^2(\phi(U),\psi(V)) &= \frac{(Cov(\phi(U),\psi(V)))^2}{Var(\phi(U))Var(\psi(V))}\\ &= \frac{(\int_{I^2}K_\theta(u,v)d\phi(u)d\psi(v))^2}{\int_{I^2}L(u,v)d\phi(u)d\phi(v)\int_{I^2}L(u,v)d\psi(u)d\psi(v)}.\\ &= \frac{\langle\phi,\psi\rangle_\theta^2}{\langle\phi,\phi\rangle_1\langle\psi,\psi\rangle_1}\end{aligned}$$

As $C_\theta(u,v)$ is symmetric in u,v, we have $Cov(\phi(U),\psi(V)) = Cov(\psi(U),\phi(V))$, so we can consider only canonical functions such that $\phi \equiv \psi$.

Let us adapt the definition of canonical functions and correlations to the continuous case. We define the classes

$$\begin{aligned} \mathcal{C}(\gamma_1,\varepsilon) &= \{\mathcal{H}_{\gamma,\varepsilon}, \gamma \in (\gamma_1, 1]\}, \\ \mathcal{C}(\gamma_1,\varepsilon)^\perp &= \{\phi \mid \langle \phi, \mathcal{H}_{\gamma,\varepsilon}\rangle_\theta = 0, \gamma \in (\gamma_1,1]\}. \end{aligned}$$

Then given $\mathcal{C}(\gamma_1,\varepsilon)$ we should seek ϕ_2 such that

$$\rho(\phi_2(U),\phi_2(V)) = \text{maximum constrained to } \phi_2 \in C(\gamma_1,\varepsilon)^\perp.$$

Next, we evaluate

$$\langle \mathcal{H}_{\gamma_1,\varepsilon}, \mathcal{H}_{\gamma_2,\varepsilon}\rangle_\theta = \int_0^1\int_0^1 K_\theta(u,v)d\mathcal{H}_{\gamma_1,\varepsilon}(u)d\mathcal{H}_{\gamma_2,\varepsilon}(v).$$

On one hand

$$\int_0^1 K_\theta(u,v)d\mathcal{H}_{\gamma_2,\varepsilon}(v) = \begin{cases} u[\gamma_2^{1-\theta} - \gamma_2(\gamma_2+\varepsilon)^{-\theta}] & \text{if } \quad 0 < u < \gamma_2, \\ \gamma[u^{1-\theta} - u(\gamma+\varepsilon)^{-\theta}] & \text{if } \gamma_2 \le u \le \gamma_2+\varepsilon, \\ 0 & \text{if } \quad 0 < \gamma_2+\varepsilon < u < 1. \end{cases}$$

On the other hand, if $\gamma_1 + \varepsilon < \gamma_2$, where ε is arbitrarily small,

$$\begin{aligned} \langle \mathcal{H}_{\gamma_1,\varepsilon}, \mathcal{H}_{\gamma_2,\varepsilon}\rangle_\theta &= \textstyle\int_0^{\gamma_2} u[\gamma_2^{1-\theta} - \gamma_2(\gamma_2+\varepsilon)^{-\theta}]d\mathcal{H}_{\gamma_1,\varepsilon}(u) \\ &+ \textstyle\int_{\gamma_2}^{\gamma_2+\varepsilon} \gamma_2[u^{1-\theta} - u(\gamma_2+\varepsilon)^{-\theta}]d\mathcal{H}_{\gamma_1,\varepsilon} \\ &+ \textstyle\int_{\gamma_2}^1 0 d\mathcal{H}_{\gamma_1,\varepsilon}(u) \\ &= 0. \end{aligned}$$

However

$$\begin{aligned} \textstyle\int_0^\gamma u[\gamma^{1-\theta} - \gamma(\gamma+\varepsilon)^{-\theta}]d\mathcal{H}_{\gamma,\varepsilon}(u) &= 0, \\ \textstyle\int_{\gamma+\varepsilon}^1 0 d\mathcal{H}_{\gamma,\varepsilon}(u) &= 0, \\ \textstyle\int_\gamma^{\gamma+\varepsilon} \gamma[u^{1-\theta} - u(\gamma+\varepsilon)^{-\theta}]d\mathcal{H}_{\gamma,\varepsilon}(u) &= \gamma(\gamma^{1-\theta} - \gamma(\gamma+\varepsilon)^{-\theta}). \end{aligned}$$

Thus

$$\begin{aligned} \langle \mathcal{H}_{\gamma,\varepsilon}, \mathcal{H}_{\gamma,\varepsilon}\rangle_\theta &= \gamma(\gamma^{1-\theta} - \gamma(\gamma+\varepsilon)^{-\theta}), \\ \langle \mathcal{H}_{\gamma,\varepsilon}, \mathcal{H}_{\gamma,\varepsilon}\rangle_1 &= \gamma(1 - \gamma(\gamma+\varepsilon)^{-1}). \end{aligned} \tag{5}$$

where $\langle \mathcal{H}_{\gamma,\varepsilon}, \mathcal{H}_{\gamma,\varepsilon}\rangle_1 = Var(\mathcal{H}_{\gamma,\varepsilon}(U))$. Therefore $\langle \mathcal{H}_{\gamma_1,\varepsilon},\ \mathcal{H}_{\gamma_2,\varepsilon}\rangle_\theta = 0$ whereas $\langle \mathcal{H}_{\gamma,\varepsilon}, \mathcal{H}_{\gamma,\varepsilon}\rangle_\theta = ||\mathcal{H}_{\gamma,\varepsilon}||_\theta^2 \neq 0$.

Theorem 2. *The set of canonical functions and canonical correlations is* $(\phi_\gamma, \lambda_\gamma)$, $0 \le \gamma \le 1$, *where* ϕ_γ *is the indicator of* γ *and* $\lambda_\gamma = \theta\gamma^{1-\theta}$.

Proof. Integrating (4) with respect to $d\mathcal{H}_{\gamma,\varepsilon}(u)$ we get $\rho(\mathcal{H}_{\gamma,\varepsilon}(U),\mathcal{H}_{\gamma,\varepsilon}(V))=\lambda_\gamma(\varepsilon)$, which tends to $\lambda_\gamma=\theta\gamma^{1-\theta}$ as $\varepsilon\to 0$. We can find the same correlation using (5).

We have $\langle\mathcal{H}_{\gamma_1,\varepsilon},\mathcal{H}_{\gamma,\varepsilon}\rangle_\theta = Cov(\mathcal{H}_{\gamma_1,\varepsilon}(U),\mathcal{H}_{\gamma,\varepsilon}(V)) = 0$ if $\gamma\neq\gamma_1$. Thus $\mathcal{H}_{\gamma,\varepsilon}\in C(\gamma_1,\varepsilon)^\perp$ and

$$\lim_{\varepsilon\to 0}\sup_{\gamma\leq\gamma_1}\rho(\mathcal{H}_{\gamma,\varepsilon}(U),\mathcal{H}_{\gamma,\varepsilon}(V))=\sup_{\gamma\leq\gamma_1}\rho(\phi_\gamma(U),\phi_\gamma(V))$$

is attained at γ_1. This maximal correlation constrained to $\gamma<\gamma_1$ is the eigenvalue $\theta\gamma_1^{1-\theta}$. Then, as $\varepsilon\to 0$, we get $C(\gamma_1)=\{\phi_\gamma|\gamma\in(\gamma_1,1]\}$ and $\phi_{\gamma'}\in C(\gamma_1)^\perp$ if $\gamma'<\gamma_1$.

To check the maximal correlation, let $g(U)$ be a function of U, where $g\in\ell^2([0,1])$ is continuous. We can suppose $g\geq 0$. Define

$$g_n(\omega)=\sum_{i=1}^n g(v_i)\mathcal{H}_{v_i,\varepsilon}(\omega),$$

where $0\leq v_1<\cdots<v_n\leq 1$. Then

$$g_n(\omega)=g(v_i)+\sum_{k<i}\frac{\varepsilon}{v_k+\varepsilon}g(v_k)\quad \text{if}\quad v_i\leq\omega\leq v_i+\varepsilon.$$

The $\mathcal{H}_{v_i,\varepsilon}$ are orthogonal with respect to $\langle\cdot,\cdot\rangle_\theta$, so $||g_n||_\theta^2=\sum_{i=1}^n g(v_i)^2||\mathcal{H}_{v_i,\varepsilon}||_\theta^2$.

As $\lambda_v(\varepsilon)=||\mathcal{H}_{v,\varepsilon}||_\theta^2/||\mathcal{H}_{v,\varepsilon}||_1^2<1$ is increasing in v, the squared correlation $\rho(g_n(U),\,g_n(V))$ is

$$\begin{aligned}\frac{||g_n||_\theta^2}{||g_n||_1^2} &= \frac{\sum_{i=1}^n g(v_i)^2||\mathcal{H}_{v_i,\varepsilon}||_\theta^2}{\sum_{i=1}^n g(v_i)^2||\mathcal{H}_{v_i,\varepsilon}||_1^2}\\ &\leq \frac{||\mathcal{H}_{v_n,\varepsilon}||_\theta^2}{||\mathcal{H}_{v_n,\varepsilon}||_1^2}\\ &\leq \lambda_\gamma(\varepsilon), \qquad \text{if}\quad v_n\leq\gamma.\end{aligned}$$

We can take now $\varepsilon_n=(n\log n)^{-2}$ and $v_k=(k/n)-\varepsilon_n$. Then, if $M=\max g(v)$, with $0\leq v\leq 1$,

$$\sum_{k<i}\frac{\varepsilon_n}{v_k+\varepsilon_n}g(v_k)<M\varepsilon_n\sum_{i=1}^n\frac{n}{i}<M\frac{(n\log n+nC)}{(n\log n)^2}\to 0$$

as $n\to\infty$. Hence $g_n(\omega)\to g(\omega)$ uniformly in ω. Thus $\rho(g_n(U),g_n(V))\to\rho(g(U),g(V))$ and the maximum correlation with the above restrictions is attained at $\rho(\phi_{\gamma_1}(U),\phi_{\gamma_1}(V))$. □

Corollary 1. *The absolute maximum correlation is the dependence parameter* θ *and the associated canonical function is* ϕ_1.

Proof. The maximum value of $\theta\gamma^{1-\theta}$ is θ and is attained at $\gamma=1$. Note that, at $x=1$, we can identify the indicator ϕ_1 with the Heaviside distribution $\mathcal{H}_1$. □

References

Cuadras, C.M. (2002a) On the covariance between functions. *Journal of Multivariate Analysis*, **81**, 19-27.

Cuadras, C.M. (2002b) Correspondence analysis and diagonal expansions in terms of distribution functions. *Journal of Statistical Planning and Inference*, **103**, 137-150.

Cuadras, C.M., Augé, J. (1981) A continuous general multivariate distribution and its properties. *Communications in Statistics-A, Theory Methods,* **10**(4), 339-353.

Cuadras, C.M., Fortiana, J. (1995) A continuous metric scaling solution for a random variable. *Journal of Multivariate Analysis,* **52**, 1–14.

Cuadras, C.M., Fortiana, J., Greenacre, M.J. (1999) Continuous Extensions of Matrix Formulations in Correspondence Analysis, with Applications to the FGM Family of Distributions. In: *Innovations in Multivariate Statistical Analysis,* R.D.H. Heijmans, D.S.G. Pollock and A. Satorra eds., Kluwer Ac. Publ., Dordrecht, 101-116.

Genest, C., Plante, J-F. (2003) On Blest's measure of rank correlation. *The Canadian Journal of Statistics*, **31**, 1-18.

Lancaster, H.O. (1969) *The Chi-squared Distribution.* Wiley, London.

Nelsen, R. B. (1999) *An Introduction to Copulas.* Springer, New York.

Ruiz-Rivas, C., Cuadras, C.M. (1988) Inference properties of a one-parameter curved exponential family of distributions with given marginals. *Journal of Multivariate Analysis*, **27**, 447-456.