Modelling extreme dependence for multivariate data

Damien Bosc [*]   Alfred Galichon [†]

EXTENDED SUMMARY

Extreme dependence between two real random variables (X, Y) is characterized by the fact that their cumulative distribution function is given by

$$F_{X,Y}(x, y) = \min(F_X(x), F_Y(y))$$

or equivalently their copula is the upper Fréchet copula

$$C(u_1, u_2) = min(u_1, u_2)$$

This form of dependence occurs when X and Y are comonotonic, i.e. when both X and Y can be written as a nondecreasing function of a third random variable Z (for instance one can choose $Z = X + Y$). As a consequence, comonotonic variables maximize covariance :

$$\mathbf{E}(XY) = \sup_{\substack{\tilde{X} \sim X \\ \tilde{Y} \sim Y}} \mathbf{E}(\tilde{X}\tilde{Y}) \tag{1}$$

When dealing with the multivariate case, where X and Y are random vectors in $\mathbf{R}^n$, there is no canonical way to generalize this notion of extreme dependence and Fréchet copula.

One possible generalization arises from the theory of optimal transport. The following optimization problem

$$\max_{\substack{\tilde{X} \sim X \\ \tilde{Y} \sim Y}} \mathbf{E} < \tilde{X}, \tilde{Y} > \tag{2}$$

is a multivariate extension of the covariance maximization problem (1) : one looks for the joint random vectors $(\tilde{X}, \tilde{Y})$ such that $\tilde{X} \sim X$, $\tilde{Y} \sim Y$ which maximizes the covariance. Under weak assumptions on the marginals, the Brenier theorem ensures that the optimal coupling (X_{opt}, Y_{opt}) satisfies the following : Y_{opt} can be written as the gradient of a convex function of X_{opt}. This is the

[*] AXA-IM and Economics Department at Ecole Polytechnique
[†] Economics Department at Ecole Polytechnique

generalization in the multivariate case of Y_{opt} being an increasing function of X_{opt}.

However this is not fully satisfying, as it only takes into account the component-wise dependence between X and Y. Indeed $< \tilde{X}, \tilde{Y} >= Tr(Cov(\tilde{X}, \tilde{Y})$ and all the information about cross-dependence between X and Y is lost.

A weaker and new way to generalize extreme dependence is to look for the couplings $(\tilde{X}, \tilde{Y})$ that yields to a *covariance matrix* $Cov(\tilde{X}, \tilde{Y}) = \mathbf{E}(\tilde{X}\tilde{Y}') = (\mathbf{E}(\tilde{X}_i\tilde{Y}_j))_{i,j}$ which would be maximal for a certain partial ordering on matrices.

More precisely, we say that a coupling (X, Y) is bigger than another coupling $(\tilde{X}, \tilde{Y})$ if their symmetric covariance matrices $\Sigma_{X,Y} = \frac{1}{2}(\mathbf{E}(XY') + \mathbf{E}(YX'))$ satisfies $\Sigma_{X,Y} - \Sigma_{\tilde{X},\tilde{Y}}$ is definite positive.

We then show that maximal couplings for this ordering induce optimal couplings in the sense of (2) replacing the scalar product $<,>$ by a semidefinite positive bilinear form. This result is the consequence of a variational characterization of the maximality of covariance matrices. As a consequence, it yields to a broader class of extremely dependent variables than the covariance maximization problem (2) : for instance a couple (X, Y) such that $Cov(X_1, Y_1)$ or $Cov(X_1, Y_2)$ is maximal would be a maximal coupling.

It has also an interpretation in terms of multivariate risk measures ; in our definition the maximally dependent couplings (X, Y) are those who minimize the maximum correlation risk measure (2) over the directions SY, S being symmetric nonnegative.

This generalizes extreme dependence and Fréchet Copula to the multivariate case and has applications in finance when high dimensionality can not be avoided.