Physica A 534 (2019) 122181

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A projection pricing model for non-Gaussian financial returns

Ana Flávia P. Rodrigues [a], Charles C. Cavalcante [b,*], Vicente L. Crisóstomo [a]

[a] Departamento de Contabilidade, Universidade Federal do Ceará, Av da Universidade, 2486, Benfica, 60020-180, Fortaleza - CE, Brazil
[b] Departamento de Engenharia de Teleinformática, Universidade Federal do Ceará, Campus do Pici, Bloco 722, CP 6005, 60440-900, Fortaleza - CE, Brazil

HIGHLIGHTS

- Portfolio design for heavy tail distributions (non-Gaussian), modeled by q-exponential models.
- Information geometrical aspects of the distributions of capital asset pricing model (CAPM).
- New measure of risk based on statistical divergence models.
- Mean–divergence model generalizes Markowitz (mean–variance) portfolio design.

ARTICLE INFO

Article history:
Received 3 April 2019
Received in revised form 26 July 2019
Available online 8 August 2019

Keywords:
CAPM
Optimal portfolio
Information geometry
Mean–divergence model

ABSTRACT

Stephen LeRoy, Jan Werner and David Luenberger have developed a geometric approach to the capital asset pricing model (CAPM) in terms of projections in a Hilbert space onto a mean–variance efficient frontier. Using this projection method, they were able to elegantly deduce a geometric interpretation of CAPM and factor asset pricing models. In this paper we extend their geometric methods to non-Euclidean divergence geometries. This extension has relevant consequences. First, it permits to deal with higher order moments of the probability distributions since general statistical divergences could encode global information about these distributions as is the case of the entropy. Secondly, orthogonal Euclidean projections and the corresponding least squares problem give place to Riemannian projections onto a possibly curved efficient frontier. Finally, our method is flexible enough to deal with huge families of probability distributions. In particular, there is no need to assume normality of the returns of the financial assets.

1. Introduction

The formulation of a non-extensive Statistical Physics by C. Tsallis [1,2] and collaborators has been developed along the last two decades in a wide range of applications to complex systems, particularly in Finance topics [3–7].

As highlighted by J. Naudts, deformed exponentials play a central role in the foundations of that Generalized Thermostatistics. Indeed, Naudts' work established deep and fruitful connections between Statistical Physics and Information Geometry [8–11]. For instance, both Rényi's and Tsallis' entropies are described by Naudts in terms of statistical divergences in the family of q-exponential distributions which includes the q-Gaussian distributions, defined in details by A. Plastino and C. Vignat [11–15].

* Corresponding author.
E-mail addresses: anaflaviapaivar@gmail.com (A.F.P. Rodrigues), charles@gtel.ufc.br (C.C. Cavalcante), vlc@ufc.br (V.L. Crisóstomo).

https://doi.org/10.1016/j.physa.2019.122181
0378-4371/

The analytic and geometric features of deformed exponentials suggest that they are well suited to model non-normally distributed returns of contingent claims. In this direction, for instance, a non-Gaussian option pricing theory has been successfully proposed in terms of diffusion processes associated to q-Gaussian distributions [6,7,16–18]. Other related developments are also summarized in [4,5].

Up to our knowledge, however, a systematic theory of portfolio optimization and asset pricing in the context of deformed exponentials has not yet been fully formalized. The classical Markowitz's mean–variance model of portfolio selection relies on the assumptions that the returns of assets are normally distributed and that the investor preferences are described by constant risk-aversion utility function.

The traditional criticism to the normality assumption in Markowitz's theory raises the need of alternative models for dealing with non-Gaussian distributions. This question has been addressed since then under different methods. In [19,20], the authors extended the Markowitz's model to the wider family of exponential distributions, replacing the mean–variance by a mean–divergence model. The idea is that Bregman divergences replace the variance as risk measures for non-Gaussian distributions, eventually encompassing information from higher order moments. On the other hand, since statistical divergences define geometric notions on the statistical manifold of exponential distributions, their method has a geometric interpretation in terms of a steepest descent by the *natural* gradient of the risk premium [21–23].

In [24], the authors propose a model of portfolio selection of financial assets that explores the non-additivity and non-normality aspects of Tsallis' Thermostatistics. More precisely, they have extended the mean–divergence model in [19,20] to the deformed exponentials families.

This paper is a natural sequel of [24] in the sense we propose here a generalization of beta (systematic risk) pricing models adapted to a mean–divergence portfolio selection [25–27]. In particular, we present an extension of the Capital Asset Pricing Model (CAPM) by Sharpe [25], Lintner [26] and Moussin [27], one of the cornerstones of the modern Finance Theory. In spite of its quite restrictive underlying assumptions, CAPM is still one of the most pervasive tools in the financial market since it establishes a simple linear relation between the risk premium for risky assets and the market risk premium. The coefficients in those linear relations are usually denominated *betas* and indicate the share of the systematic risk (the market risk, say) as a component of the variance of the excess return of a particular financial asset. The generalized model we propose is flexible enough to be applied for financial returns with deformed exponential distributions. Our method relies on a geometric approach to the classical mean–variance analysis developed by S. LeRoy and J. Werner [28] and D. Luenberger [29], see also [30] and [31].

This paper is structured as follows. In Section 2 we define the geometric setting of the space of contingent claims $\mathcal{M}$ and the subspace of traded financial assets $\mathcal{M}$ in terms of statistical manifolds of probability distributions as the manifold of ϕ-deformed exponentials and some geometric features of that manifold are summarized in this section. Section 3 is devoted to the description of the notation used in the derivations of the paper and to highlight the main results of the work. Discussion on the model of pricing based on the projection and its relation with the mean–divergence optimization model is presented in Section 4. We deduce in Section 5 an expression of a minimum divergence portfolio in the efficient frontier. As in the classical beta pricing models, the proportions of market portfolio and risk-free assets in this optimal portfolio are dictated by a linear regression coefficient but in our case we consider the Riemann curvature to encompass the third and fourth order moments of the distribution of returns, generalizing the variance which is the case for flat (Euclidean) spaces. Section 6 is dedicated to one of our main results, namely the generalized CAPM and the efficient market portfolio. Finally, we state our conclusions and discuss on some new directions of research in Section 7.

2. The space of financial assets

Following [28,29,31], one models the set $\mathcal{M}$ spanned by traded financial assets as a subspace of a Hilbert space $\mathcal{H}$ of contingent claims. More precisely, every point in $\mathcal{M}$ corresponds to the payoff z of a contingent claim at a fixed time, say $t = 1$, that is, a random variable

$$z = z(s),$$

where s are the states of the world with probability distribution specified by some density $p(s; \boldsymbol{\vartheta})$. Here, $\boldsymbol{\vartheta}$ is the distribution parameter of a family of probability distributions whose densities define a n-dimensional statistical manifold

$$\mathcal{S} = \{p(s, \boldsymbol{\vartheta}) : \boldsymbol{\vartheta} \in U \subset \mathbb{R}^n\},$$

where $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_n)$ takes values in some open subset U of the n-dimensional Euclidean space $\mathbb{R}^n$. In the sequel, the fact that the random variable z depends on the states of the world s whose probability distribution is $p(s, \boldsymbol{\vartheta})$ will be summarized by the notation $z \sim p(s, \boldsymbol{\vartheta})$. Sometimes we also express this by writing $z(s, \boldsymbol{\vartheta})$.

Example 1. Suppose that $\mathcal{M}$ is spanned by finitely many assets and that the space of states of the world is also finite-dimensional. Then each $z \in \mathcal{M}$ is determined by their possible payoffs in distinct (and also finitely many) states of the world, say, $\{s_1, \ldots, s_N\}$. In other terms, each $z \in \mathcal{M}$ is described by a N-dimensional vector

$$z = (z(s_1), \ldots, z(s_N)).$$

The (discrete) probability distribution of the states of the world is denoted by $p_i = p(s_i; \boldsymbol{\vartheta})$, $i = 1, \ldots, N$, where $\boldsymbol{\vartheta}$ indicates the parameters of this distribution. For the sake of simplicity, in what follows we will restrict ourselves to a finite-dimensional asset span $\mathcal{M}$. The case of infinite-dimensional Hilbert spaces can be handled with some notational and technical adjustments.

In sum, a point $z \in \mathcal{M}$ corresponds to the possible payoffs of a given asset under the distinct states of the world. The probability distribution of these states is given by a probability density in a statistical manifold.

2.1. Geometry of statistical divergences

In [28] and [29], S. LeRoy, J. Werner and D. Luenberger have developed a geometric approach to the capital asset pricing model (CAPM) in terms of a Hilbert space geometry of projections onto a mean–variance efficient frontier. Using this projection method, they easily deduce an elegant geometric interpretation of CAPM and factor pricing models.

In what follows, we extend their geometric methods to divergence geometries in $\mathcal{M}$ more general than the Euclidean geometry induced from the Hilbert space norm. Denoting by $\langle \cdot, \cdot \rangle$ the inner product in $\mathcal{M}$ induced from $\mathcal{H}$, we can assume that D is a Bregman divergence of the form

$$D(z|w) = K(z) - K(w) - \langle \nabla K(w), z - w \rangle \tag{1}$$

for some convex function $K : \mathcal{M} \to \mathbb{R}$. Here ∇K is the Fréchet differential of K on $\mathcal{M}$, which corresponds to the usual gradient in the case when $\mathcal{M}$ is finite-dimensional and $\langle \cdot, \cdot \rangle$ is the Euclidean inner product.

A trivial yet fundamental example of statistical divergence is the square of the Euclidean L^2-norm in $\mathcal{H}$ restricted to $\mathcal{M}$, that is

$$D_{\text{euc}}(z|w) = \frac{1}{2}|z - w|^2, \quad z, w \in \mathcal{M}. \tag{2}$$

In the sequel we are going to consider more general examples, not necessarily quadratic. For instance, we may fix the Kullback–Leibler divergence [21]

$$D_{KL}(z(\cdot, \boldsymbol{\vartheta})|w(\cdot, \boldsymbol{\vartheta}')) = \int p(s, \boldsymbol{\vartheta}) \log \left(\frac{p(s, \boldsymbol{\vartheta})}{p(s, \boldsymbol{\vartheta}')} \right) \mathrm{d}s. \tag{3}$$

The following example is naturally linked to the statistical manifolds of deformed exponentials with probability densities of the form

$$p(s, \boldsymbol{\vartheta}) = \exp_\phi(\langle T(s), \boldsymbol{\vartheta} \rangle - K(\boldsymbol{\vartheta}))\, p_0(\boldsymbol{\vartheta}), \quad \boldsymbol{\vartheta} \in \mathbb{R}^n, \tag{4}$$

where T is a given smooth function of the random variable $z(s) \sim p(s, \boldsymbol{\vartheta})$ and K is the cumulant function. Here, p_0 is a fixed reference density and $\exp_\phi$ is the ϕ-exponential defined as the inverse function of the ϕ-logarithm [8,9]

$$\log_\phi(t) = \int_1^t \frac{1}{\phi(s)} \mathrm{d}s,$$

where $\phi : (0, +\infty) \to (0, +\infty)$ is a strictly positive, nondecreasing and continuous real function. A particular case of this *deformed* exponential is given by the q-exponential function

$$\exp_q(t) = (1 + (1 - q)t)^{\frac{1}{1-q}}$$

with $q > 0$, $q \neq 1$, what corresponds to set $\phi(t) = t^q$, Hence, the q-logarithm is defined by

$$\log_q(t) = \int_1^t \frac{1}{s} \mathrm{d}s = \frac{1}{1-q}(t^{1-q} - 1).$$

Setting $\phi(t) = t$ one gets the family of exponential distributions, in particular multivariate Gaussian distributions. For any choice of ϕ as above, we set the statistical divergence given by the relative ϕ-entropy

$$D_\phi(z(\cdot, \boldsymbol{\vartheta})|w(\cdot, \boldsymbol{\vartheta}')) = \mathbb{E}_{\hat{p}(\cdot, \boldsymbol{\vartheta})} \left[\log_\phi p(\cdot, \boldsymbol{\vartheta})/p_0 - \log_\phi p(\cdot, \boldsymbol{\vartheta}')/p_0 \right], \tag{5}$$

where $\hat{p}(s, \boldsymbol{\vartheta})$ is the escort distribution [8,9] given by

$$\hat{p}(s, \boldsymbol{\vartheta}) = \frac{1}{h(\boldsymbol{\vartheta})} \psi \left(\langle T(s), \boldsymbol{\vartheta} \rangle - K(\boldsymbol{\vartheta}) \right) p_0(s)$$

with

$$h(\boldsymbol{\vartheta}) = \int \phi \left(p(s, \boldsymbol{\vartheta})/p_0(s) \right) \mathrm{d}s$$

and

$$\psi(t) = \phi \left(\exp_\phi(t) \right).$$

Note that D_{KL} in (3) corresponds to D_ϕ in (5) for the particular choice of $\phi(t) = t$.

In all of those examples, namely (2), (3) and (5), the cumulant function K defines a Bregman divergence as in (1) where the probability distributions of the random variables $z(s)$ and $w(s)$ are respectively given by the densities $p(s, \boldsymbol{\vartheta})$ and $p(s, \boldsymbol{\vartheta}')$. By a slight abuse of notation, we will sometimes write $K(z)$ instead of $K(\boldsymbol{\vartheta})$ to indicate that K is evaluated at the random variable $z \in \mathcal{M}$ whose probability density is $p(\cdot, \boldsymbol{\vartheta})$.

The Hessian of the cumulant function K defines a Riemannian metric g in $\mathcal{M}$ whose contravariant version g^* is the Hessian of the dual cumulant function K^*, that is, the Legendre transform of K defined by

$$K^*(z(s, \boldsymbol{\eta})) = \max_{\boldsymbol{\vartheta}} \left(\langle \boldsymbol{\vartheta}, \boldsymbol{\eta} \rangle - K(z(s, \boldsymbol{\vartheta})) \right) .$$

In the case of ϕ-deformed exponentials (4)–(5) the dual function K^* is given by the relative negative ϕ-entropy [32]

$$K^*(z(\cdot\,, \boldsymbol{\eta})) = \mathbb{E}_\phi[\log_\phi p(\cdot\,, \boldsymbol{\vartheta})/p_0(\cdot)]$$

where $\mathbb{E}_\phi[\cdot]$ is the expectation taken over the deformed exponential ϕ (also called ϕ-expectation) [24] and $\boldsymbol{\eta}$ is the dual affine coordinate defined by

$$\boldsymbol{\eta} = \nabla K(z(\cdot\,, \boldsymbol{\vartheta})).$$

In the particular example of q-exponential distributions one has

$$K^*(z(\cdot\,, \boldsymbol{\eta})) = \frac{1}{1-q} \left(\frac{1}{h(\boldsymbol{\vartheta}) - 1} \right)$$

with

$$h(\boldsymbol{\vartheta}) = \int \left(p(s, \boldsymbol{\vartheta})/p_0(s) \right)^q p_0(s)\, \mathrm{d}s,$$

where $\boldsymbol{\vartheta}$ is the statistical parameter of the distribution $p(s, \boldsymbol{\vartheta})$ of the payoff $z = z(s, \boldsymbol{\vartheta})$. In the particular case of $\phi(t) = t$ that corresponds to the exponential family of distributions (that includes Gaussian distributions) we have

$$g|_z = \mathrm{var}[z],$$

the variance taken with respect to the probability density

$$p(s, \boldsymbol{\vartheta}) = \exp(\langle T(s), \boldsymbol{\vartheta} \rangle - K(s, \boldsymbol{\vartheta}))\, p_0(\boldsymbol{\vartheta})$$

when $z \sim p(s, \boldsymbol{\vartheta})$. As we will discuss in the sequel, Gaussian and, more generally, exponential distributions correspond to the choice of the Euclidean divergence (2) in the geometric projection approach to the asset pricing, [28,29].

In general, the metrics g and g^* define a dually flat structure with affine connections whose geodesics are Euclidean lines in terms of the coordinates $\boldsymbol{\vartheta}$, $\boldsymbol{\eta}$ in $\mathcal{M}$. We refer the reader to [21,33] for a comprehensive account of those concepts in terms of Information Geometry. One of the fundamental results in Information Geometry is the Pythagorean Theorem that can be stated in the following form

Theorem 1 (*Theorem 1.2 and Theorem 1.3, [21]*)**.** *Given $o, z, w \in \mathcal{M}$ such that the dual affine geodesic connecting z and w is orthogonal to the affine geodesic connecting w and o, the following generalized Pythagorean relation holds*

$$D(z|o) = D(w|o) + D(z|w). \tag{6}$$

Similarly, if the affine geodesic connecting z and w is orthogonal to the dual affine geodesic connecting w and o we have the dual relation

$$D^*(z|o) = D^*(w|o) + D^*(z|w), \tag{7}$$

where D^* is the dual Bregman divergence

$$D^*(z|w) = K^*(z) - K^*(w) - \langle \nabla K^*(w), z - w \rangle. \tag{8}$$

This geometric relation will be used in what follows to deduce a general single factor model for asset pricing in $\mathcal{M}$. Other versions of the previous Theorem can be found in [15,34] where additional requirements about the pairs of points o, z and w are assumed/imposed. Such versions are more suitable for our purposes in this work so will use the one provided in [15].

The rationale of the Pythagorean Theorem in Finance applications can be found, for example, in [35] and references therein.

3. Notation and main results

A statistical divergence in $\mathcal{M}$ as in (1) defines a Riemannian metric given by the Hessian of the convex function K, that is,

$$g|_z = \nabla^2 K(z). \tag{9}$$

This means that for each $z \in \mathcal{M}$ we have a inner product $g|_z$, that is, a symmetric bilinear positive definite form acting on vectors tangent to z. Since $\mathcal{M}$ is (at least locally) represented by a linear space we can think of a vector ξ tangent to $\mathcal{M}$ at z as a displacement vector of the form $w - z$ for some $w \in \mathcal{M}$. Hence, the action of $g|_z$ can be expressed in local coordinates as

$$g|_z(\xi, \xi) = g_{ij}(z)\xi^i \xi^j,$$

where $(g_{ij}(z))$ is a positive definite symmetric matrix whose components depend smoothly on the coordinates of $z \in \mathcal{M}$. Here, ξ^i represent the coordinates of the tangent vector ξ. Note that

$$g_{ij}(z) = \left.\frac{\partial^2 K}{\partial z^i \partial z^j}\right|_z.$$

Fixed this notation, we consider an *expectation kernel*, that is, an asset in $\mathcal{M}$ that yields the expected payoffs of the assets in $\mathcal{M}$. More precisely

$$g|_z(k_{\mathrm{e}}, z) = \mathbb{E}[z] \tag{10}$$

for any $z \in \mathcal{M}$. We also fix a *pricing kernel* k_{q} as an asset in $\mathcal{M}$ that gives the price of any contingent claim $z \in \mathcal{M}$ as the expected discounted payoff

$$g|_z(k_{\mathrm{q}}, z) = \mathbb{E}[\mu z] = \mathrm{q}(z), \tag{11}$$

where μ is a stochastic discount factor. Here $\mathrm{q} : \mathcal{M} \to \mathbb{R}$ is the price functional, that is, the present value of the expected returns of the asset, discounted at rate μ. Since the expectation is not necessarily taken with respect to risk-neutral probabilities, μ is a risk-adjusted discount rate, possibly distinct from the risk-free return rate [28,36].

In geometric terms, the kernels k_{e} and k_{q} are tangent vector fields in $\mathcal{M}$ that represent the linear functionals $\mathbb{E}$ and q using the Riemannian metric g in $\mathcal{M}$ as in (10) and (11), respectively. Denote by $\mathcal{E}$ the subspace in $\mathcal{M}$ spanned by k_{e} and k_{q}. The *projection* $z^{\mathcal{E}}$ of $z \in \mathcal{M}$ onto $\mathcal{E}$ is defined by

$$D(z|z^{\mathcal{E}}) = \min_{w \in \mathcal{E}} D(z|w).$$

It follows from the generalized Pythagorean Theorem for divergences (Theorem 1) that fixed a reference point $o \in \mathcal{M}$ one has

$$D(z|o) = D(z^{\mathcal{E}}|o) + D(z|z^{\mathcal{E}}), \tag{12}$$

for $z \in \mathcal{M}$. One of the requirements to Eq. (12) hold is that one of the pairs $(z, z^{\mathcal{E}})$ or $(z^{\mathcal{E}}, o)$ must lie on a geodesic which is affine as a function of expectation while the other pair lies on a geodesic which is affine as a function of the parameters.

In the particular case of the divergence given by the Euclidean L^2-norm in $\mathcal{M}$

$$D_{\mathrm{euc}}(z|w) = \frac{1}{2}|z - w|^2 \tag{13}$$

expression (12) reduces to the Euclidean decomposition

$$|z|^2 = \mathbb{E}[z]^2 + \mathrm{var}[z], \tag{14}$$

where

$$\mathrm{var}[z] = \mathbb{E}[(z - \mathbb{E}[z])^2]$$

is the variance, the classical risk measure in Portfolio Theory [36,37]. In this case, the orthogonal projection onto $\mathcal{E}$ is the solution of the following *least squares* optimization problem: given $z \in \mathcal{M}$, to find $z^{\mathcal{E}}$ in $\mathcal{E}$ such that

$$|z - z^{\mathcal{E}}|^2 = \min_{w \in \mathcal{E}} |z - w|^2 \tag{15}$$

the solution of which minimizes the variance amongst all the payoffs z whose orthogonal projection in $z^{\mathcal{E}}$.

Motivated by the analogy between (12) and (14), we define the projection

$$\mathcal{P}(z) = D(z|z^{\mathcal{E}})$$

as a geometrically natural surrogate of the variance in the setting of general statistical divergences. This replacement has relevant consequences. First, it permits to deal with higher order moments of the probability distributions since general statistical divergences could encode global information about these distributions as is the case of the entropy. Secondly, orthogonal Euclidean projections and the corresponding least squares problem give place to Riemannian projections onto principal curves and surfaces. Finally, this proposal is flexible enough to deal with huge families of probability distributions going far beyond normality assumptions about returns of financial assets.

Now we are able to state our main results. The following theorem states that the two reference assets k_e and k_q determine the efficient frontier for portfolios of assets in $\mathcal{M}$. Hence, it generalizes the projection pricing method to the context of non-Euclidean statistical divergences. Indeed we have

Theorem 2. *Let $\mathcal{E}$ be the subspace in $\mathcal{M}$ spanned by the expectation kernel k_e and the pricing kernel k_q. Given $z \in \mathcal{M}$ we have*

$$\mathbb{E}[z^{\mathcal{E}}] = \mathbb{E}[z]$$

and

$$\mathcal{P}[z^{\mathcal{E}}] \leq \mathcal{P}[z]$$

where $z^{\mathcal{E}}$ is the projection of z onto $\mathcal{E}$ in the sense that

$$D(z|z^{\mathcal{E}}) = \min_{w \in \mathcal{E}} D(z|w). \tag{16}$$

Denote by R_e and R_q the returns of k_e and k_q, respectively. Hence, the counterpart of the pricing equations in [29] in the context of the mean–divergence efficient frontier $\mathcal{E}$ reads as

Theorem 3. *The minimum divergence portfolio in $\mathcal{M}$ is given by*

$$z = R_\mathrm{e} + (1 - \beta)(R_\mathrm{q} - R_\mathrm{e}) \tag{17}$$

where

$$\beta = -\frac{g(R_\mathrm{q} - R_\mathrm{e}, R_\mathrm{e})}{g(R_\mathrm{q} - R_\mathrm{e}, R_\mathrm{q} - R_\mathrm{e})} \tag{18}$$

with

$$g = \nabla^2 K(z).$$

It is worth to note that it is convenient for practical purposes to work with the expansion of the Riemannian metric around a fixed reference point $o \in \mathcal{M}$ as

$$g|_z \sim \nabla^2 K(o) + o(|z|^2), \tag{19}$$

where quadratic terms are determined in terms of the Riemann curvature of the Riemannian manifold $(\mathcal{M}, g)$, see [38]. Then, one can replace g by the Hessian matrix $\nabla^2 K(o)$ ignoring higher order corrections.

In the case when there is a risk-free asset $\mathbf{1}$ in $\mathcal{M}$ with return $\bar{R}$ we obtain from Theorem 3 a generalized CAPM expression of the form

$$\mathbb{E}[z] = \bar{R} + \beta(\mathbb{E}[R_\mathrm{q}] - \bar{R}) \tag{20}$$

as a consequence of our projection method based on minimizing a given statistical divergence. The expression (20) extends the classical CAPM formula to the setting of non-normal distributions. Besides that, this general formula is not necessarily deduced from a least squares minimization since we are considering divergences other than the Euclidean one.

For instance, in the particular case when we suppose that the returns of traded assets are distributed accordingly a q-Gaussian distributions it holds that

$$g|_z = \nabla^2 K(z(\cdot, \boldsymbol{\vartheta})) = \Sigma_q$$

for every $z \in \mathcal{M}$, where the q-variance matrix Σ_q is defined in Section 7. Hence, we get the following consequence of Theorem 3.

Corollary 1. *Suppose that the traded financial assets in $\mathcal{M}$ are distributed according to a q-Gaussian distribution. Hence the minimum divergence portfolio is given by*

$$z = R_\mathrm{e} + (1 - \beta)(R_\mathrm{q} - R_\mathrm{e})$$

where

$$\beta = -\frac{g(R_\mathrm{q} - R_\mathrm{e}, R_\mathrm{e})}{g(R_\mathrm{q} - R_\mathrm{e}, R_\mathrm{q} - R_\mathrm{e})}$$

with

$$g = \Sigma_q,$$

being the q-variance matrix defined as

$$\Sigma_q = \gamma_q C_{q,n}^{1-q} |\Sigma|^{\frac{1-q}{2}} \Sigma \tag{21}$$

with

$$\gamma_q = \frac{1}{2}\left((n+4) - (n+2)q \right), \tag{22}$$

where n is the number of considered assets in the portfolio and

$$C_{q,n} = \begin{cases} \frac{\Gamma(\frac{1}{q-1} - \frac{n}{2})\sqrt{\pi}}{\Gamma(\frac{1}{q-1})} \left(\frac{1}{q-1} \right)^{\frac{n}{2}} \left((n+4) - (n+2)q\right)\right)^{\frac{n}{2}}, & \text{for} \quad 1 < q < \frac{n+4}{n+2}, \\ \frac{\Gamma(\frac{2-q}{1-q})\sqrt{\pi}}{\Gamma(\frac{2-q}{q-1} + \frac{n}{2})} \left(\frac{1}{1-q} \right)^{\frac{n}{2}} \left((n+4) - (n+2)q\right)\right)^{\frac{n}{2}}, & \text{for} \quad q < 1. \end{cases} \tag{23}$$

Here Σ is the variance–covariance matrix of the returns on the assets and $|\Sigma|$ is the determinant of Σ. We refer the reader to [12] for further details in q-multivariate Gaussian distributions.

Now, we describe a couple of other important facts that stem from the definition of the mean–divergence efficient frontier. In Section 5.1 we prove that beta pricing equations similar to (17)–(18) are still valid if one replaces the returns of the expectation and pricing kernels by the returns of two assets in the mean–divergence efficient frontier that are orthogonal with respect to g. This is the case of assets with zero correlation in the classical setting [28,29].

Finally, we observe that (20) becomes an exact counterpart of the CAPM equation if we could replace R_q by the return of an efficient market. This is the content of Theorem 4 in Section 6 where we present a generalized CAPM equation based on the maximization of a utility function. For that, we assume that the utility function describes the preferences of a risk averse agent that is strictly decreasing with respect to the risk measure. For such utility functions, it is possible to prove that the market equilibrium portfolio lies on the mean–divergence efficient frontier.

4. Projection pricing and mean–divergence frontier

In this section, we sketch the proof of Theorem 2.

Recall that $\mathcal{E}$ is the (one or two) dimensional subspace in $\mathcal{M}$ spanned by given payoff vectors k_e and k_q in $\mathcal{M}$. More precisely, the vector fields k_e and k_q define a (one or two dimensional) distribution and $\mathcal{E}$ is a fixed integral leaf of this distribution. Given $z \in \mathcal{M}$ we suppose that there exists $z^{\mathcal{E}} \in \mathcal{E}$ such that

$$D(z|z^{\mathcal{E}}) = \min_{w \in \mathcal{E}} D(z|w). \tag{24}$$

Fixed an arbitrary point $o \in \mathcal{M}$ it follows from Theorem 1, and also considering that $o \in \mathcal{E}$ [34], that

$$D(z|o) = D(z^{\mathcal{E}}|o) + D(z|z^{\mathcal{E}}). \tag{25}$$

At this point, it is worth to highlight the formal resemblance with the usual Pythagorean expression that takes place in the case of the Euclidean L^2-divergence (associated with the Hilbert metric)

$$D_{\text{euc}}(z|w) = \frac{1}{2}|z - w|^2, \quad z, w \in \mathcal{M}.$$

In this case, (25) is nothing but the ordinary least squares decomposition

$$|z|^2 = \mathbb{E}[z]^2 + \text{var}[z], \tag{26}$$

with $\mathbb{E}[z] = \langle z, \mathbf{1} \rangle$, where $\mathbf{1}$ indicates a risk-free asset whose payoff is 1 in every state of the world. Recall that the variance

$$\text{var}[z] = \mathbb{E}[(z - \mathbb{E}[z])^2]$$

is the classical risk measure in Portfolio Theory [36,37]. Moreover, up to a constant factor, it is the risk premium for risky assets under the assumption of normally distributed returns. As we mentioned before, comparing (25) and (26) suggests to adopt a general definition of risk premium as

$$\mathcal{P}(z) = D(z|z^{\mathcal{E}}) \tag{27}$$

in the context of non-Euclidean divergences. In this way, (24) and (25) imply that

$$\mathcal{P}(z^{\mathcal{E}}) \leq \mathcal{P}(z) \tag{28}$$

with equality only and only if $z = z^{\mathcal{E}}$. We denote $z^{\mathcal{E}} = \pi_{\mathcal{E}}(z)$.

We now deduce an infinitesimal version of the condition (24) in the case when D is a Bregman divergence with cumulant function K as in (1). Fixed $z \in \mathcal{M}$ and a curve $w(t)$ in $\mathcal{E}$ with $w(0) = z^{\mathcal{E}}$ we have

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} D(z|w(t)) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \left(K(z) - K(w(t)) - \langle \nabla K(w(t)), z - w(t) \rangle \right) \\ &= -\langle \nabla K(z^{\mathcal{E}}), w'(0) \rangle + \langle \nabla K(z^{\mathcal{E}}), w'(0) \rangle - w'(0)^{\top} \nabla^2 K(z^{\mathcal{E}})(z - z^{\mathcal{E}}), \end{aligned}$$

where $a^{\top}$ denotes the transpose of a matrix a. We conclude that

$$w'(0)^{\top} \nabla^2 K(z^{\mathcal{E}})(z - z^{\mathcal{E}}) = 0 \tag{29}$$

for an arbitrary vector $w'(0)$ tangent to $\mathcal{E}$ at $z^{\mathcal{E}}$. Therefore, denoting by ε the asset

$$\varepsilon = z - z^{\mathcal{E}} \tag{30}$$

and denoting

$$g|_{z^{\mathcal{E}}} = \nabla^2 K(z^{\mathcal{E}}), \tag{31}$$

a inner product in the tangent space $T_{z^{\mathcal{E}}}\mathcal{E}$ due to the convexity of K, we conclude that ε is perpendicular to $T_{z^{\mathcal{E}}}\mathcal{E}$. Since $\mathcal{E}$ is indeed a vector space we note that $T_{z^{\mathcal{E}}}\mathcal{E} = \mathcal{E}$. Hence, we conclude that ε is perpendicular to $\mathcal{E}$ at the point $z^{\mathcal{E}}$ with respect to the inner product $g|_{z^{\mathcal{E}}}$. In particular, it follows that

$$0 = g|_{z^{\mathcal{E}}}(\varepsilon, k_{\mathrm{e}}) = g|_{z^{\mathcal{E}}}(z, k_{\mathrm{e}}) - g|_{z^{\mathcal{E}}}(z^{\mathcal{E}}, k_{\mathrm{e}})$$

and

$$0 = g|_{z^{\mathcal{E}}}(\varepsilon, k_{\mathrm{q}}) = g|_{z^{\mathcal{E}}}(z, k_{\mathrm{q}}) - g|_{z^{\mathcal{E}}}(z^{\mathcal{E}}, k_{\mathrm{q}}).$$

Therefore

$$g|_{z^{\mathcal{E}}}(z, k_{\mathrm{e}}) = g|_{z^{\mathcal{E}}}(z^{\mathcal{E}}, k_{\mathrm{e}}) \tag{32}$$

and

$$g|_{z^{\mathcal{E}}}(z, k_{\mathrm{q}}) = g|_{z^{\mathcal{E}}}(z^{\mathcal{E}}, k_{\mathrm{q}}) \tag{33}$$

for any z with projection $z^{\mathcal{E}}$, that is, z of the form $z^{\mathcal{E}} + t\varepsilon$, $t \in \mathbb{R}$. This means that we may extend the Riemannian metric $g|_{z^{\mathcal{E}}}$ in points of $\mathcal{E}$ to a Riemannian metric in the whole space $\mathcal{M}$ simply declaring that it is invariant with respect to translations in directions ε perpendicular to $\mathcal{E}$. In other terms, the Riemannian metric g in $\mathcal{M}$ is fixed in such a way that $\mathcal{M}$ becomes the Riemannian product $\mathcal{M} = \mathcal{E} \times \mathcal{F}$, where $\mathcal{F}$ is ruled by Euclidean lines of the form $z^{\mathcal{E}} + t\varepsilon$, $t \in \mathbb{R}$, with $z^{\mathcal{E}} \in \mathcal{E}$ and ε perpendicular to the tangent space $T_{z^{\mathcal{E}}}\mathcal{E}$ at $z^{\mathcal{E}}$ with respect to $g|_{z^{\mathcal{E}}}$.

This choice of a product Riemannian metric permits to interpret expressions (32) and (33) in terms of expected values and prices of assets. Indeed, the vector field k_{e} tangent to $\mathcal{E}$ can be interpreted as an *expectation kernel* in the sense that

$$g|_{z^{\mathcal{E}}}(k_{\mathrm{e}}, z) = \mathbb{E}[z] = \sum_{i=1}^{N} p_i z(s_i). \tag{34}$$

In the same way one interprets the vector field k_{q} tangent to $\mathcal{E}$ as the *pricing kernel* that must satisfy

$$g|_{z^{\mathcal{E}}}(k_{\mathrm{q}}, z) = \mathbb{E}[\mu z] = \sum_{i=1}^{N} p_i \mu(s_i) z(s_i) = \mathrm{q}[z]. \tag{35}$$

Here, μ is a stochastic discount factor. Hence,

$$k_{\mathrm{e}}|_{z^{\mathcal{E}}} = \left(g|_{z^{\mathcal{E}}} \right)^{-1} \left(p_1, \ldots, p_N \right) \tag{36}$$

and

$$k_{\mathrm{q}}|_{z^{\mathcal{E}}} = \left(g|_{z^{\mathcal{E}}} \right)^{-1} \left(\mu(s_1)p_1, \ldots, \mu(s_N)p_N \right), \tag{37}$$

where $\mu(s_i)$ is the discount rate in the scenario relative to the state of the world s_i, $i = 1, \ldots, N$. This interpretation of k_{e} and k_{q} implies that we can rewrite (32) and (33) respectively as

$$\begin{aligned} \mathbb{E}[z] &= \mathbb{E}[z^{\mathcal{E}}], \\ \mathrm{q}[z] &= \mathrm{q}[z^{\mathcal{E}}]. \end{aligned} \tag{38}$$

In particular $\mathbb{E}[\varepsilon] = 0$. In sum, payoffs in $\mathcal{M}$ have the same expected values and prices than their orthogonal projections on $\mathcal{E}$ with respect to the Riemannian metric g.

It also follows from the fact that g is a Riemannian product metric that the Riemannian curvature of $\mathcal{M}$ is determined by the Gaussian curvature $\mathcal{K}$ of $\mathcal{E}$ in the case when k_{e} and k_{q} are linearly independent and $\mathcal{M}$ is a two-dimensional vector space. In this case, considering for instance a local coordinate system given by the principal directions of $\nabla^2 K(z^{\mathcal{E}})$ one has

$$\mathcal{K}|_{z^{\mathcal{E}}} = -\frac{1}{2\sqrt{\lambda_1\lambda_2}}\left(\partial_1\left(\frac{\partial_1\lambda_2}{\sqrt{\lambda_1\lambda_2}}\right)+\partial_2\left(\frac{\partial_2\lambda_1}{\sqrt{\lambda_1\lambda_2}}\right)\right), \tag{39}$$

where $\lambda_1, \lambda_2 > 0$ are the eigenvalues of $g|_{z^{\mathcal{E}}} = \nabla^2 K(z^{\mathcal{E}})$. Note that the Gaussian curvature involves third and fourth moments of the distributions $p(s, \boldsymbol{\vartheta})\,ds$. In particular, $\mathcal{E}$ and $\mathcal{M}$ become flat spaces if we consider statistical divergences that depend only on second order moments as is the case of the Euclidean divergence (2). Indeed, if K is quadratic, as in the example of the Euclidean divergence D_{euc}, we have

$$\nabla^2 K(z^{\mathcal{E}}) = \nabla^2 K(o),$$

where $o \in \mathcal{M}$ is an arbitrarily fixed reference point. In this case,

$$g = \nabla^2 K(o)$$

and $\mathcal{M}$ is a flat Riemannian manifold. This is exactly the context of the classical Euclidean theory. In the general case we have an expansion of the form

$$\nabla^2 K(z^{\mathcal{E}}) \simeq \nabla^2 K(o) + o(|z|^2),$$

where the quadratic remainder encodes the Riemannian curvature of $\mathcal{M}$ and its covariant derivatives. In statistical terms, these curvature terms can be associated to the contribution of higher moments of the underlying probability distributions.

5. Minimum divergence portfolio

In this section, we present the proof of Theorem 3.

We have proved in Section 4 that $\mathcal{E} = \text{span}\{k_q, k_e\}$ is the mean–divergence frontier in $\mathcal{M}$. Now we address the problem of minimizing the risk measure

$$\mathcal{P}(z) = D(o|z)$$

among points in $z \in \mathcal{E}$ only, that is,

$$\min_{z\in\mathcal{E}} D(o|z), \tag{40}$$

where $o \in \mathcal{M}$ is an arbitrarily fixed reference point. Any point $z \in \mathcal{E}$ is of the form

$$z = ak_q + bk_e,$$

for some $a, b \in \mathbb{R}$. The price of this portfolio is

$$\mathsf{q}(z) = \mathbb{E}[\mu z] = a\mathbb{E}[\mu k_q] + b\mathbb{E}[\mu k_e] = a\mathsf{q}(k_q) + b\mathsf{q}(k_e) \tag{41}$$

Fixing the constraint that the price of the portfolio is $\mathsf{q}(z) = 1$, we denote

$$\beta = a\mathsf{q}(k_q)$$

and therefore

$$1 - \beta = b\mathsf{q}(k_e).$$

Therefore the portfolios with unit price are parameterized by

$$z = \beta\frac{k_q}{\mathsf{q}(k_q)} + (1-\beta)\frac{k_e}{\mathsf{q}(k_e)} = \beta R_q + (1-\beta)R_e = R_e + \beta(R_q - R_e) \tag{42}$$

with $\beta \in \mathbb{R}$. Here, R_q and R_e are the returns of k_q and k_e, respectively. Then, minimizing the risk premium among payoffs in $\mathcal{E}$ with unit price turns out to be equivalent to the one-dimensional minimization problem

$$\min_{\beta} D(o|R_e + \beta(R_q - R_e)),$$

whose first order necessary condition is

$$0 = \frac{d}{d\beta}D(o|R_e + \beta(R_q - R_e)) = -(R_q - R_e)^\top \nabla^2 K(R_e + \beta(R_q - R_e))(R_e + \beta(R_q - R_e)).$$

We conclude that the optimal portfolio with unit price is determined by

$$\beta = -\frac{(R_q - R_e)^\top \nabla^2 K(R_e + \beta(R_q - R_e))R_e}{(R_q - R_e)^\top \nabla^2 K(R_e + \beta(R_q - R_e))(R_q - R_e)}. \tag{43}$$

Considering the approximation

$$\nabla^2 K(R_e + \beta(R_q - R_e)) \simeq \nabla^2 K(o),$$

we fix an *approximate* value of β that determines the choice of optimal portfolio $z_0 \in \mathcal{E}$ with unit price by

$$\beta_0 := -\frac{(R_q - R_e)^\top \nabla^2 K(o) R_e}{(R_q - R_e)^\top \nabla^2 K(o)(R_q - R_e)} \quad (44)$$

Note that the expected return of this portfolio is

$$\mathbb{E}[z_0] = \mathbb{E}[R_e] + \beta_0\, \mathbb{E}[R_q - R_e] \quad (45)$$

We have in the case when the risk-free asset **1** with riskless return $\bar{R}$ is an element in $\mathcal{M}$ that

$$k_e = \mathbf{1}$$

and $R_e = \mathbb{E}(\mathbf{1}) = \bar{R}$. Hence, in this case

$$\mathbb{E}[z_0] = \bar{R} + \beta_0\,(\mathbb{E}[R_q] - \bar{R}), \quad (46)$$

which is similar to the classical beta pricing equation obtained when we consider the Euclidean divergence (2) for which $\nabla^2 K(o)$ equals the variance.

5.1. Generalized beta pricing

From now on, we assume that $\mathcal{E}$ is two-dimensional. Hence, it is convenient to rewrite the pricing equations above using two linearly independent assets other than k_e and k_q. We fix such assets, say k_λ and k_μ, with respective returns

$$r_\lambda = R_e + \lambda(R_q - R_e)$$

and

$$r_\mu = R_e + \mu(R_q - R_e)$$

in such a way that

$$g|_o(r_\lambda, r_\mu) = 0. \quad (47)$$

Hence, μ is given by

$$\mu = -\frac{g|_o(R_e, R_e) + \lambda g|_o(R_q - R_e, R_e)}{g|_o(R_q - R_e, R_e) + \lambda g|_o(R_q - R_e, R_q - R_e)} \quad (48)$$

Note that μ is well-defined if and only if $\lambda \neq \beta_0$ in (44), that is, if k_λ is not the (approximate) minimum divergence portfolio in $\mathcal{E}$.

Given an asset $z \in \mathcal{M}$ with unit price we have the decomposition

$$z = z^{\mathcal{E}} + \varepsilon$$

where

$$z^{\mathcal{E}} = ak_\lambda + bk_\mu$$

with ε perpendicular to $\mathcal{E}$ at $z^{\mathcal{E}}$ and $\mathbb{E}[\varepsilon] = 0$. It follows that

$$\begin{aligned} \mathbb{E}[z] &= a\mathbb{E}[k_\lambda] + b\mathbb{E}[k_\mu] = a\mathrm{q}(k_\lambda)\mathbb{E}[r_\lambda] + b\mathrm{q}(k_\mu)\mathbb{E}[r_\mu] \\ &=: \mathbb{E}[r_\mu] + \beta(\mathbb{E}[r_\lambda] - \mathbb{E}[r_\mu]) \end{aligned}$$

with $\beta = a\mathrm{q}(k_\lambda)$. Since z has unit price the return r of z is given by

$$r = z = a\mathrm{q}(k_\lambda) r_\lambda + b\mathrm{q}(k_\mu) r_\mu + \varepsilon = r_\mu + \beta(r_\lambda - r_\mu) + \varepsilon$$

from which it follows that

$$\begin{aligned} g|_{z^{\mathcal{E}}}(r, r_\lambda) &= g|_{z^{\mathcal{E}}}(r_\mu, r_\lambda) + \beta g|_{z^{\mathcal{E}}}(r_\lambda - r_\mu, r_\lambda) + g|_{z^{\mathcal{E}}}(\varepsilon, r_\lambda) \\ &= g|_o(r_\mu, r_\lambda) + o(|z|^2) + \beta g|_{z^{\mathcal{E}}}(r_\lambda - r_\mu, r_\lambda) \\ &= \beta g|_{z^{\mathcal{E}}}(r_\lambda - r_\mu, r_\lambda) + o(|z|^2). \end{aligned}$$

We conclude that

$$\beta = \frac{g|_{z^{\mathcal{E}}}(r, r_\lambda)}{g|_{z^{\mathcal{E}}}(r_\lambda - r_\mu, r_\lambda)} + o(|z|^2) = \frac{g|_o(r, r_\lambda)}{g|_o(r_\lambda - r_\mu, r_\lambda)} + o(|z|^2) = \frac{g|_o(r, r_\lambda)}{g|_o(r_\lambda, r_\lambda)} + o(|z|^2),$$

where $g|_o = \nabla^2 K(o)$. In sum, we have obtained a generalized beta pricing equation

$$\mathbb{E}[z] = \mathbb{E}[r_\mu] + \beta(\mathbb{E}[r_\lambda] - \mathbb{E}[r_\mu]) \tag{49}$$

for assets in $z \in \mathcal{M}$, where the generalized beta coefficient is approximated (up to quadratic remainder terms) by

$$\beta = \frac{g|_o(r, r_\lambda)}{g|_o(r_\lambda, r_\lambda)}. \tag{50}$$

If the risk-free asset $\mathbf{1}$ with return $\bar{R}$ lies in the space of traded contingent claims $\mathcal{M}$, we fix $r_\mu = \mathbf{1}$. With this choice, (49) reduces to

$$\mathbb{E}[z] = \bar{R} + \beta(\mathbb{E}[r_\lambda] - \bar{R}), \tag{51}$$

a generalized beta pricing equation written in terms of an asset k_λ instead of the return of the pricing kernel k_q as in (46).

5.2. Optimal portfolio for q-exponentials

In this section, we focus on the particular example of the statistical manifold of ϕ-exponential probability densities defined in (4), recalling the results from [24].

Recall that setting $\phi(t) = t$ in (4) one gets the family of exponential distributions, in particular multivariate Gaussian distributions. For this family, R. Nock, B. Magdalou, E. Bryis and F. Nielsen [19,20] represented the key concepts of Portfolio Selection theory in terms of the cumulant function and the associated Bregman divergence. More precisely, they proved that for CARA utility functions the certainty equivalent and risk premium of risky assets are respectively given by

$$\mathcal{C} = \frac{1}{a}\left(K(z) - K(w) \right)$$

and

$$\mathcal{P} = \frac{1}{a} D[z|w],$$

where $a > 0$ is a risk-aversion parameter. Hence they extended the classical mean–variance portfolio selection to a general mean–divergence model for which an optimal allocation $\boldsymbol{\alpha}$ is a solution of the minimization problem

$$\min_{\boldsymbol{\alpha}} \left(\langle \nabla K(z(s, \boldsymbol{\vartheta} - a\boldsymbol{\alpha})), \boldsymbol{\alpha} \rangle + \frac{1}{a} D(z(s, \boldsymbol{\vartheta})|z(s, \boldsymbol{\vartheta} - a\boldsymbol{\alpha})) \right).$$

In the particular case of Gaussian distributed returns, one recovers the classical Markowitz's optimal portfolio allocation vector

$$\boldsymbol{\alpha} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T \Sigma^{-1}\mathbf{1}},$$

where Σ is the variance–covariance matrix of the returns on the assets.

In [24], the authors extended this mean–divergence approach to ϕ-exponential distributions, in particular to q-exponential distributions. They proved that the optimal portfolio for their extended mean–divergence model is given in terms of the cumulant function by

$$\boldsymbol{\alpha} = \frac{\nabla^2 K(z(s, \boldsymbol{\vartheta}))^{-1}\mathbf{1}}{\mathbf{1}^T \nabla^2 K(z(s, \boldsymbol{\vartheta}))^{-1}\mathbf{1}}. \tag{52}$$

Note that the Hessian $\nabla^2 K$ of the cumulant (convex) function K is positive-definite and plays the role of the variance–covariance matrix in the Gaussian case. In the particular case of q-Gaussian distributions [12], the optimal allocation portfolio is given by

$$\boldsymbol{\alpha} = \frac{\Sigma_q^{-1}\mathbf{1}}{\mathbf{1}^T \Sigma_q^{-1}\mathbf{1}} \tag{53}$$

where Σ_q is defined according to (21). It is evident that one reobtains the Markowitz's portfolio as $q \to 1$ in (53).

In view of (52), the authors have elaborated in [24] a steepest descent algorithm by the natural (Riemannian) gradient of the risk premium. Some empirical support to the proposed method is provided by comparing the cumulated returns and the evolution of the divergence for optimal portfolios according to the mean–divergence model and the classical one by Markowitz. The numerical evaluations in [24] show the proposed method yields better tracking of deep changes in the stock market, such as the ones present in economic crisis.

6. Efficient market portfolio and generalized CAPM

As in the classical CAPM, we can take r_λ as the market return r_{m} since it is possible to prove under some assumptions that r_{m} is in the mean–divergence efficient frontier $\mathcal{E}$. In this case, both (49) and (51) define a generalized security market line [28,39].

Suppose that every agent in the market has consumption preferences given by a time-separable utility function of the form

$$u(c_0, c_1) = u_0(c_0) + u_1(\mathbb{E}[c_1], g|_{c_1}(c_1, c_1)) \tag{54}$$

where u_1 is strictly decreasing with respect to the second variable. Here c_0 is the agent's consumption plan at time $t = 0$ and $c_1 = c_1(s)$ is a random variable in $\mathcal{M}$ that describes the consumption plan of the agent at time $t = 1$.

The optimal agent's consumption plan is a solution of the constrained intertemporal optimization problem

$$\max_{c_0, c_1, \boldsymbol{\alpha}} u(c_0, c_1)$$

subject to the constraints

$$c_0 \le w_0 - \boldsymbol{\alpha} \cdot \mu z,$$
$$c_1 \le w_1 + \boldsymbol{\alpha} \cdot z$$

where z is a portfolio of risky traded assets in $\mathcal{M}$, $\boldsymbol{\alpha}$ is the portfolio allocation vector and w_0 and w_1 are, respectively, the agent's endowments at time $t = 0$ and $t = 1$. If we suppose for the sake of simplicity that we have an interior optimal solution then the first-order condition reads as

$$\mu z = \frac{\partial_{c_1} u}{\partial_{c_0} u} z,$$

where the ratio on the right hand side is the marginal rate of substitution for the utility function u, see [40]. Taking expected values on both sides one gets

$$\mathsf{q}(z) = \mathbb{E}\left[\frac{\partial_{c_1} u}{\partial_{c_0} u} z\right]$$

for the optimal agent's consumption plan c_1. Suppose without loss of generality that c_1 is tradable, that is, it is asset in $\mathcal{M}$. We then prove that c_1 lies in the mean–divergence frontier $\mathcal{E}$. We consider the orthogonal decomposition

$$c_1 = \pi_{\mathcal{E}}(c_1) + c_1^{\perp}$$

where

$$g|_{c_1}(c_1^{\perp}, \mathcal{E}) = 0.$$

Then we define an alternative consumption plan by $\widetilde{c}_1 = \pi_{\mathcal{E}}(c_1)$. Suppose by contradiction that $c_1^{\perp} > 0$. We conclude that

$$\widetilde{c}_1 - w_1 < c_1 - w_1.$$

Moreover since $\mathsf{q}(c_1^{\perp}) = \mathbb{E}(c_1^{\perp}) = 0$ we have

$$\mathsf{q}(\widetilde{c}_1 - w_1) = \mathsf{q}(c_1 - w_1)$$

and

$$\mathbb{E}[\widetilde{c}_1 - w_1] = \mathbb{E}[c_1 - w_1].$$

We also have $\pi_{\mathcal{E}}(c_1) = \pi_{\mathcal{E}}(\widetilde{c}_1)$ and

$$g|_{\widetilde{c}_1}(\widetilde{c}_1, c_1) \le g|_{c_1}(c_1, c_1).$$

Since the agent's preferences are described by an utility function that is strictly increasing with respect to the risk measure (the second variable), we conclude that $\widetilde{c}_1$ is strictly preferred to c_1. This contradicts the optimality of the consumption plan c_1. From this contradiction, we conclude that $c_1 \in \mathcal{E}$ for every agent. Since the market payoff z_{m} is by definition the sum over agents of the tradable components of agents' consumption plans, the market return lies on the mean–divergence frontier as well.

In view of the above, we now deduce a generalized CAPM equation.

Theorem 4. *The equilibrium prices for efficient assets z with returns r in a market with agents' preferences described by an utility function of the form (54) are given by*

$$\mathbb{E}[r] - \mathbb{E}[r_\mu] = \widetilde{\beta}(\mathbb{E}[r_{\mathsf{m}}] - \mathbb{E}[r_\mu]), \tag{55}$$

where r_m *is the return of the market portfolio and*

$$\widetilde{\beta} = \frac{g(r, r_m)}{g(r_m, r_m)}. \tag{56}$$

Proof. We have proved that the tradable component of the optimal consumption plan lies in the mean–divergence efficient frontier $\mathcal{E}$. Then for each agent, labeled by $k = 1, \ldots, M$, we have

$$c_{0,(k)} \leq w_{0,(k)} - q(\boldsymbol{\alpha}_{(k)}) \cdot z - b_{i,(k)},$$

where the portfolio $\boldsymbol{\alpha} \cdot z_{(k)}$ lies in the mean–divergence efficient frontier and satisfies $g(\boldsymbol{\alpha}_{(k)} \cdot z, r_\mu) = 0$ for every $k = 1, \ldots, M$. Here, $b_{i,(k)}$ is the share invested in the asset with return r_μ (that could represent the risk-free asset, in the case when it is available in the market $\mathcal{M}$). Hence, at time $t = 1$ this inequality becomes

$$c_{1,(k)} \leq w_{1,(k)} + \boldsymbol{\alpha}_{(k)} \cdot z + b_{i,(k)} r_\mu.$$

Since $q(\boldsymbol{\alpha}_{(k)} \cdot z) = \boldsymbol{\alpha}_{(k)} \cdot \mu z$ we have

$$\mathbb{E}[c_{1,(k)}] = \mathbb{E}[c_{1,(k)}] = (w_{0,(k)} - c_{0,(k)})\mathbb{E}[r_\mu] + \boldsymbol{\alpha}_{(k)} \cdot (\mathbb{E}[z] - \mu \mathbb{E}[r_\mu])$$

and using that $g(\boldsymbol{\alpha}_{(k)} \cdot z, r_\mu) = 0$ we have

$$g(c_{1,(k)}, c_{1,(k)}) = g(\boldsymbol{\alpha}_{(k)} \cdot z, \boldsymbol{\alpha}_{(k)} \cdot z) + b_k^2 g(r_\mu, r_\mu).$$

Differentiating $u_{(k)}$, the utility function for the preferences of the kth agent, at an equilibrium portfolio with respect to the allocation parameter $\boldsymbol{\alpha}$ one obtains the vector equation

$$\partial_1 u_{(k)}(\mathbb{E}[z_i] - \mu_i \mathbb{E}[r_\mu]) + 2\partial_2 u_{(k)} \boldsymbol{\alpha}^j_{(k)} g(z_i, z_j) = 0$$

from what follows that the optimal allocation for each agent is given by

$$\boldsymbol{\alpha}^j_{(k)} = -\frac{\partial_1 u_{(k)}}{\partial_2 u_{(k)}} g(z_i, z_j)^{-1}(\mathbb{E}[z_j] - \mu_i \mathbb{E}[r_\mu])$$

Summing up on $k = 1, \ldots, M$, one gets

$$\gamma^{-1} g(z_i, z_j)^{-1}(\mathbb{E}[z_j] - \mu_j \mathbb{E}[r_\mu]) = 1$$

where

$$\gamma = -\left(\sum_{k=1}^{M} \frac{\partial_1 u_{(k)}}{\partial_2 u_{(k)}} \right)^{-1}$$

Denoting $g_{ij} = g(z_i, z_j)$ one concludes that the market equilibrium price for each asset z_i is given by

$$\mu_i = \frac{1}{\mathbb{E}[r_\mu]} \left(\mathbb{E}[z_i] - \gamma g_{ij} 1^j \right)$$

Hence we have

$$\mathbb{E}[r_i] = \frac{1}{\mu_i}\mathbb{E}[z_i] = \mathbb{E}[r_\mu] + \frac{\gamma}{\mu_i} g_{ij} 1^j = \mathbb{E}[r_\mu] + \gamma g \left(\frac{z_i}{\mu_i}, \sum_i z_i \right) = \mathbb{E}[r_\mu] + \gamma g \left(\frac{z_i}{\mu_i}, z_m \right).$$

We conclude that

$$\mathbb{E}[r_i] - \mathbb{E}[r_\mu] = \gamma \mu_m g(r_i, r_m)$$

where $\mu_m = z_m / r_m$ is the value of the market payoff at $t = 0$. Denoting

$$\widetilde{\beta}_i = \frac{g(r_i, r_m)}{g(r_m, r_m)} \tag{57}$$

one obtains

$$\mathbb{E}[r_i] - \mathbb{E}[r_\mu] = \gamma \widetilde{\beta}_i \mu_m g(r_m, r_m).$$

In particular,

$$\mathbb{E}[r_m] - \mathbb{E}[r_\mu] = \gamma \mu_m g(r_m, r_m).$$

Therefore

$$\mathbb{E}[r_i] - \mathbb{E}[r_\mu] = \widetilde{\beta}_i(\mathbb{E}[r_m] - \mathbb{E}[r_\mu]). \tag{58}$$

This finishes the proof. □

7. Conclusions and future works

In this paper we have proposed a generalization of the capital asset pricing model. For this sake, we defined expectation and price kernels in terms of a statistical divergence, particularly the Bregman divergence, in a manifold which contains the assets and used the Riemannian metric derived from the used divergence. This idea is motivated by the consideration of the divergence as an alternative risk measure, instead of the using the variance which is the natural metric for normal distributions.

Hence, we deduced an expression of a minimum divergence portfolio in the efficient frontier. As in the classical beta pricing models, the proportions of market portfolio and risk-free assets in this optimal portfolio are dictated by a linear regression coefficient. However, in our model, we take into account the Riemannian metric in the manifold $\mathcal{M}$, defined by the traded assets, which is given by the Hessian of the cumulant function K of the deformed exponential probability density. This approach makes possible to generalize the Gaussian distribution cases, where the flat metric considers only the variance (second order moments) of the portfolio. In our general approach, the Riemann curvature of $\mathcal{M}$ encodes third and fourth order moments of the distribution of returns.

We are currently obtaining further developments for applications of the theoretical models elaborated in this work. One of those ongoing projects are related to estimation techniques of the generalized beta factors, specially useful for valuation models in Corporate Finance.

Acknowledgments

The authors would like to thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior — Brazil (CAPES) - Finance Code 001 and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil (Procs. 309472/2017-2 and 408609/2016-8) for partial funding of this research. The authors also thanks the anonymous reviewers who provided very important comments, specially on the requirements of Theorems 1 and 2.

References

[1] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys. 52 (1–2) (1988) 479–487.
[2] C. Tsallis, D.J. Bukman, Anomalous diffusion in the presence of external forces: Exact time-dependent solutions and their thermostatistical basis, Phys. Rev. E 54 (1996) R2197.
[3] C. Tsallis, Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World, Springer-Verlag, New York, 2009.
[4] C. Tsallis, C. Anteneodo, L. Borland, R. Osorio, Nonextensive statistical mechanics and economics, Physica A 324 (1–2) (2003) 89–100.
[5] C. Tsallis, Economics and finance: q-statistical stylized features galore, Entropy 19 (9) (2017) 457–475.
[6] L. Borland, A theory of non-Gaussian option pricing, Quant. Finance 2 (6) (2002) 415–431.
[7] B. Trivellato, Deformed exponentials and applications to finance, Entropy 15 (9) (2013) 3471–3489.
[8] J. Naudts, Generalised Thermostatistics, Springer, 2011.
[9] J. Naudts, Estimators, escort probabilities and ϕ-exponential families in statistical physics, J. Inequal. Pure Appl. Math. 5 (4) (2004) 102.
[10] J. Naudts, On the maximum entropy principle in non-extensive thermostatistics, 2004, arXiv:cond-mat/0405508v1.
[11] J. Naudts, Generalised exponential families and associated entropy functions, Entropy 10 (3) (2008) 131–149.
[12] A.R. Plastino, A. Plastino, Non-extensive statistical mechanics and generalized Fokker-Planck equation, Physica A 222 (1–4) (1995) 347–354.
[13] C. Vignat, A. Plastino, Central limit theorem and deformed exponentials, J. Phys. A 40 (45) (2007).
[14] J. Naudts, Deformed exponentials and logarithms in generalized thermostatistics, Physica A 316 (1–4) (2002) 323–334.
[15] J. Naudts, B. Anthonis, The exponential family in abstract information theory, in: F. Nielsen, F. Barbaresco (Eds.), Geometric Science of Information, Springer, Berlin, Heidelberg, 2013, pp. 265–272.
[16] L. Borland, J.-P. Bouchaud, A non-Gaussian option pricing model with skew, Quant. Finance 4 (5) (2004) 499–514.
[17] L. Borland, Option pricing formulas based on a non-Gaussian stock price model, Phys. Rev. Lett. 89 (9) (2002).
[18] E. Moretto, S. Pasquali, B. Trivellato, Option pricing under deformed Gaussian distributions, Physica A 446 (2016) 246–263.
[19] R. Nock, B. Magdalou, E. Briys, F. Nielsen, On tracking portfolios with certainty equivalents on a generalization of Markowitz model: The fool, the wise and the adaptive, in: Proceedings of the 28th International Conference on International Conference on Machine Learning, ICML'11, 2011, pp. 73–80.
[20] R. Nock, B. Magdalou, E. Briys, F. Nielsen, Mining matrix data with Bregman matrix divergences for portfolio selection, in: F. Nielsen, R. Bhatia (Eds.), Matrix Information Geometry, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 373–402.
[21] S.-I. Amari, Information Geometry and Its Applications, Springer, 2016.
[22] S.-I. Amari, Natural gradient works efficiently in learning, Neural Comput. 10 (2) (1998) 251–276.
[23] S. Amari, A. Cichocki, Information geometry of divergence functions, Bull. Polish Acad. Sci. Tech. Sci. 58 (1) (2010) 183–195.
[24] A.F.P. Rodrigues, I.M. Guerreiro, C.C. Cavalcante, Deformed exponentials and portfolio selection, Internat. J. Modern Phys. C 29 (3) (2018) http://dx.doi.org/10.1142/S0129183118500298.
[25] W.F. Sharpe, CaPital asset prices: A theory of market equilibrium under conditions of risk, J. Finance 19 (3) (1964) 425–442.
[26] J. Lintner, The aggregation of investor's diverse judgments and preferences in purely competitive security markets, J. Financ. Quant. Anal. 4 (4) (1969) 347–400.
[27] J. Moussin, Equilibrium in a capital asset market, Econometrica 34 (4) (1966) 768–783.
[28] S. LeRoy, J. Werner, Principles of Financial Economics, Cambridge University Press, 2014.
[29] D.G. Luenberger, Projection pricing, J. Optim. Theory Appl. 109 (1) (2001) 1–25.
[30] C. Fontana, M. Schweizer, Simplified mean-variance portfolio optimisation, Math. Financ. Econ. 6 (2) (2012) 125–152.
[31] D. Duffie, Dynamic Asset Pricing Theory, Princeton University Press, 2001.
[32] S.-i. Amari, A. Ohara, Geometry of q-exponential family of probability distributions, Entropy 13 (6) (2011) 1170–1185.
[33] N. Ay, J. Jost, H.V. Le, L. Schwachhöfer, Information Geometry, Springer, 2017.
[34] B. Anthonis, Extension of Information Geometry for Modelling Non-Statistical Systems (Ph.D. thesis), Universiteit Antwerpen, 2014.

[35] A.F.P. Rodrigues, C.C. Cavalcante, Principal curves for statistical divergences and an application to finance, Entropy 20 (5) (2018) http://dx.doi.org/10.3390/e20050333.
[36] K. Back, Asset Pricing and Portfolio Choice Theory, Oxford University Press, 2010.
[37] H. Markowitz, Portfolio selection, J. Finance 7 (1) (1952) 77–91.
[38] J. Jost, Riemannian Geometry and Geometric Analysis, Springer, 2011.
[39] D.G. Luenberger, Investment Science, Oxford University Press, 1998.
[40] A. Mas-Collel, M. Whinston, J. Green, Microeconomic Theory, Oxford University Press, 1995.