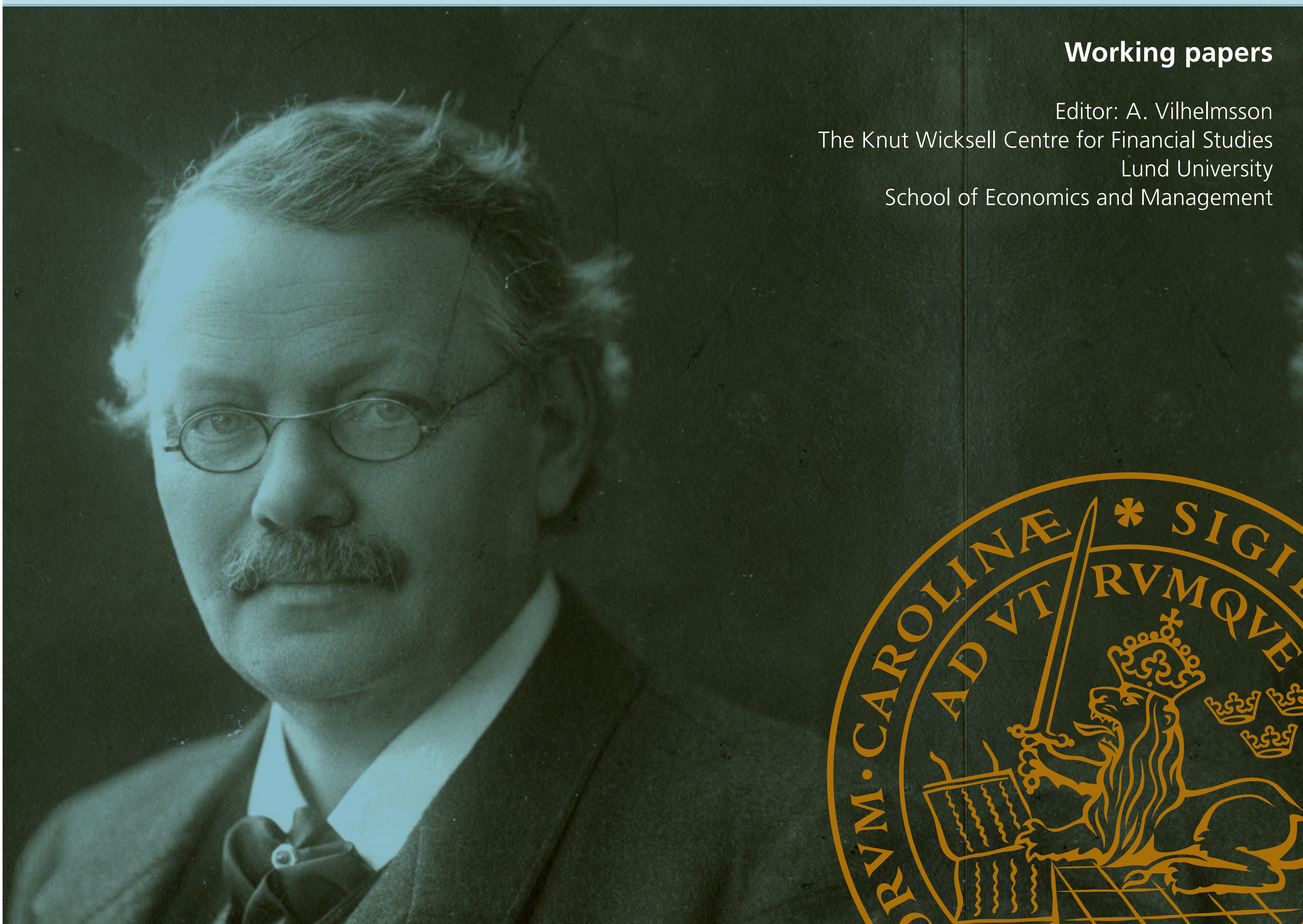
The Geometry of Risk Adjustments

HANS-PETER BERMIN AND MAGNUS HOLM

KNUT WICKSELL WORKING PAPER 2021:2



Working papers

Editor: A. Vilhelmsson
The Knut Wicksell Centre for Financial Studies
Lund University
School of Economics and Management

The Geometry of Risk Adjustments

Hans-Peter Bermin* Magnus Holm†

December 7, 2021

Abstract

In this paper we present a geometric approach to portfolio theory, with the aim to explain the geometrical principles behind risk adjusted returns; in particular Jensen's alpha. We find that while the alpha/beta approach has severe limitations (especially in higher dimensions), only minor conceptual modifications are needed to complete the picture. However, these modifications (e.g. using risk adjusted Sharpe ratios rather than returns) can only be appreciated once a full geometric approach to portfolio theory is developed. We further show that, in a complete market, the so called market price of risk vector is identical to the growth optimal Kelly vector, albeit expressed in coordinates of a different basis. For trading strategies collinear to the growth optimal Kelly vector, we formalise a notion of relative value trading based on the risk adjusted Sharpe ratio. As an application we show that a derivative having a risk adjusted Sharpe ratio of zero has a corresponding price given by the the minimal martingale measure.

Keywords: Jensen's alpha, Kelly criterion, market price of risk, option pricing, geometry

*Corresponding author, Knut Wicksell Centre for Financial Studies, Lund University, Sweden and Hilbert Group. E-mail: hans-peter@hilbertcapital.com.

†Hilbert Group.

1 Introduction

In this paper we present a geometric approach to portfolio theory, in part based on the framework outlined in Bermin and Holm (2021b). More precisely, we consider an opportunity set consisting of N primary assets and a numéraire asset, such that a self-financing trading strategy can, for a fixed point in time, be seen as an element in $\mathbb{R}^N$. This vector space is further endowed with a natural inner product, through the instantaneous covariance matrix of logarithmic excess returns, which thus forms a Hilbert space. As shown in Bermin and Holm (2021b), the instantaneous rate of excess portfolio return can then be represented as the inner product of the corresponding trading strategy and the growth optimal Kelly trading strategy. It is this unique feature that allows us to formulate a geometric approach to portfolio theory by means of a single vector (i.e. the growth optimal Kelly vector) and the inner product in our Hilbert space.

Since the growth optimal Kelly vector plays a central role to this study, we pay homage to the original contributors Kelly (1956) and Latané (1959). While early promoters exist, see for instance Hakansson and Ziemba (1995), Thorp (2011), and the references therein, the Kelly theory has nonetheless been subject to severe criticism over the years, see Ziemba (2014) for a historical recount. Recent studies, like Platen (2006), Bermin and Holm (2021b), take a more neutral approach when arguing that an investor who, for a fixed level of volatility, prefers higher rate of excess return to lower chooses to allocate the wealth proportional to the Kelly criterion. Furthermore, any such trading strategies have maximal instantaneous Sharpe ratio, see Sharpe (1966, 1994), and thus lie on the efficient (local) frontier in the sense of Markowitz (1952) and Tobin (1958). The reason being that trading strategies instantaneously uncorrelated with the growth optimal Kelly strategy have a zero rate of excess return, yet (in general) a positive volatility. Another interesting aspect of the growth optimal Kelly strategy is that the reciprocal of the portfolio value can be seen as an admissible stochastic discount factor, Long (1990). We refine this result by showing that the corresponding market price of risk vector is identical to the growth optimal Kelly vector, albeit expressed in coordinates of a different basis, when the market is complete. Hence, an immediate consequence of our geometrical approach is that we strengthen the connection between these, sometimes, separate fields of research. It also provides new means to use portfolio theory in order to value derivatives in incomplete markets.

The main motivation, however, for writing this paper is to explain and clarify the geometrical principles behind risk adjusted returns; in particular Jensen's alpha as introduced in Jensen (1968). In doing so, it becomes apparent that one must also explain the meaning of the beta parameter. In this paper we let the reference asset, for both alpha and beta, be an arbitrary portfolio and not necessarily the unobservable market portfolio. The first observation we make is that given any two trading strategies, representing an investor's portfolio and the reference asset, we can achieve any targeted alpha and beta values by appropriately applying leverage to each trading strategy. From this, one cannot but conclude that the concept of risk adjusted returns is not well captured by the alpha and beta parameters. In fact, we argue that neither a higher alpha, for a fixed beta, nor a lower beta, for a fixed alpha, is strictly better for an investor. By studying the geometry of alpha and beta we show that alpha describes the excess return of a particular portfolio, formed from the investor's portfolio in such a way that it is locally uncorrelated with the reference asset. This particular portfolio consists of being long the investor's portfolio and adding beta percent of a short position in the reference asset. Forming this portfolio, however, does not tell us how best to trade in order to reach the maximal instantaneous Sharpe ratio, or equivalently to be on the efficient (local) frontier. Nor does it tell us what (logarithmic) excess return can be achieved for a given level of risk. We circumvent the problems with alpha and beta by applying the risk adjustment to the Sharpe ratio, rather than to the excess return itself. One benefit of taking this approach is that the magnitude of the Sharpe ratio is invariant with respect to leverage. Overall, we find that while the alpha/beta approach has severe limitations (especially in higher dimensions), only minor conceptual modifications are needed to complete the picture. However, these minor modifications (e.g. using risk adjusted Sharpe ratios rather than risk adjusted returns) can only be appreciated once a full geometric approach to portfolio theory is developed.

In addition, we derive a number of intermediate results that are of interest by themselves. We show that the growth optimal Kelly vector on a subspace equals the orthogonal projection of the growth optimal Kelly vector onto that subspace. We also show that the length of any growth optimal Kelly vector equals its instantaneous Sharpe ratio. A financial interpretation, of these two results, is that the maximal Sharpe ratio decreases as we reduce the opportunity set. We further show that the instantaneous correlation between an arbitrary trading strategy and its corresponding growth optimal Kelly strategy can be expressed as the ratio between their Sharpe ratios. Additionally, we derive a general bound for the correlation between two arbitrary trading strategies in terms of their Sharpe ratios and the Sharpe ratio of the corresponding growth optimal Kelly strategy. By analyzing the level sets of various financial quantities we also find that points in the mean-variance space cannot, in general, be associated with a unique trading strategy. Only the points on the efficient frontier (that is those with maximal Sharpe ratio) can uniquely be identified. For such trading strategies, collinear to the growth optimal Kelly vector, we formalise the notion of relative value trading that is implicit in Platen (2006) and Bermin and Holm (2021a). This allows us to explicitly quantify the additional return an investor can obtain for a fixed level of risk. Thereafter, we apply geometric principles to investigate derivative pricing and introduce the concept of pricing by means of No Added Relative Value (NARV, for short). We say that this concept applies to a given asset, relative an initial portfolio, when no value can be added by augmenting the initial opportunity set with the given asset. Using geometric principles we show that the NARV price of a derivative is defined such that its risk adjusted Sharpe ratio equals zero and that this price further corresponds to the no-arbitrage price of the, so called, minimal martingale measure of Föllmer and Schweizer (1991); a result first derived in Bermin and Holm (2021a), albeit with much different methods. Finally, we show how to extend the geometric approach such that risk can be measured against an arbitrary asset, different from the numéraire, as described in Bermin and Holm (2021b).

In order to derive our result we use tensor analysis. While this is not a standard approach in the financial literature, it greatly simplifies the notation and geometrical understanding, compared to a formalism based on matrix algebra. Hopefully, the readers agree with us once passed the initial hurdle.

The paper is organized as follows. In section two we briefly recap the framework laid out in Bermin and Holm (2021b). In section three we introduce basic notations from geometric algebra, after which we establish the portfolio framework in a number of subsections. In section four we investigate Jensen's alpha; starting with the simple case of how to best trade in two assets and following up with the general case of how to best trade in two opportunity sets. Section five deals with relative value trading and its connection to derivative pricing, while in section six we briefly explain how to measure risk against an asset different from the numéraire.

2 Basic Portfolio Theory

We consider a capital market consisting of a number of primary assets $(P_0, P_1, \ldots, P_N)$ expressed in some common numéraire unit, say US dollar. An asset related to a dividend paying stock is seen as a fund with the dividends re-invested. All assets are assumed to be positive adapted continuous processes living on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = \{\mathcal{F}(t) : t \geq 0\}$ is a right-continuous increasing family of σ-algebras such that $\mathcal{F}(0)$ contains all the $\mathbb{P}$-null sets of $\mathcal{F}$. As usual we think of the filtration $\mathbb{F}$ as the carrier of information. We further let P_0 be the numéraire asset of the economy, describing how the value of the numéraire unit changes over time, and introduce the relative prices $P_{0|n} = P_n/P_0$.

An investor can trade in the assets and throughout this paper we assume that there are no transaction fees, that short-selling is allowed, that trading takes place continuously in time, and that trading activity does not impact the asset prices. We define a trading strategy as an $\mathbb{F}$-predictable vector process $w = (w^1, \ldots, w^N)'$, representing the proportion of wealth invested in each asset, and we let X_w denote the corresponding portfolio. In order to analyze the performance of the numéraire based wealth process $X_{0|w} = X_w/P_0$ we impose the restriction that, when re-balancing the portfolio, money can neither be injected nor withdrawn. Such trading strategies are said to be self-financing and satisfy

$$\frac{dX_{0|w}(t)}{X_{0|w}(t)} = \sum_{n=1}^{N} w^n(t) \frac{dP_{0|n}(t)}{P_{0|n}(t)}. \tag{1}$$

Note that since $X_{0|\mathbf{0}}$ is constant, $\mathbf{0} = (0, \ldots, 0)'$, the trading strategy $X_{\mathbf{0}}$ can always be identified with the market numéraire asset P_0 for purposes when the proportionality constant plays no role. More generally, an investor can equally compare the performance of his trading strategy with an arbitrary reference strategy. We use the notation $X_{u|w} = X_w/X_u$ to refer to the ratio of portfolios using the trading strategies w and the reference strategy u, respectively. Since $X_{u|w}$ is independent of the original numéraire unit we may view X_u as the risk-free asset in the normalized capital market $(P_{u|0}, P_{u|1}, \ldots, P_{u|N})$, where $P_{u|n} = P_n/X_u$. In this setup the self-financing condition reads

$$\frac{dX_{u|w}(t)}{X_{u|w}(t)} = w^0(t) \frac{dP_{u|0}(t)}{P_{u|0}(t)} + \sum_{n=1}^{N} w^n(t) \frac{dP_{u|n}(t)}{P_{u|n}(t)}, \quad w^0(t) = 1 - \sum_{n=1}^{N} w^n(t). \tag{2}$$

The instantaneous rate of return of the trading strategy w, in excess of the reference strategy u, can therefore be expressed in terms of the instantaneous rate of excess return of the numéraire asset and the risky assets. We write this as

$$b_{u|w}(t) = w^0(t)b_{u|0}(t) + \sum_{n=1}^{N} w^n(t)b_{u|n}(t), \tag{3}$$

such that $b_{u|\mathbf{0}} = b_{u|0}$ and $b_{u|e_n} = b_{u|n}$ for $e_n = (0, \ldots, 0, 1, 0, \ldots, 0)'$ being the n'th coordinate vector corresponding to the investable assets. Note that when $u = \mathbf{0}$ and the numéraire asset is taken to be locally risk free (i.e. a bank account) we measure the rate of excess return relative the interest rate. We further define the instantaneous covariance process (by means of the quadratic covariation process, see Karatzas and Shreve (1988)) and the corresponding instantaneous variance process

$$V_{u|v,w}(t) = \frac{d}{dt}[\log X_{u|v}, \log X_{u|w}](t), \quad \sigma^2_{u|w}(t) = V_{u|w,w}(t), \tag{4}$$

while the instantaneous correlation process $\rho_{u|v,w}$ is implicitly defined by

$$V_{u|v,w}(t) = \sigma_{u|v}(t)\rho_{u|v,w}(t)\sigma_{u|w}(t). \tag{5}$$

We also let $\mu_{u|w}$ denote the rate of logarithmic excess return and claim that a simple application of Itô's formula yields

$$\mu_{u|w}(t) = b_{u|w}(t) - \frac{1}{2}\sigma^2_{u|w}(t). \tag{6}$$

Additionally, we follow Bermin and Holm (2021b) and introduce a few more important concepts. First, we define the generalized instantaneous Sharpe ratio

$$s_{u|w}(t) = \frac{b_{u|w}(t)}{\sigma_{u|w}(t)}, \tag{7}$$

Second we define the relative leverage risk processes $k_{u|w}$ and the relative drawdown process

$R_{u|w}$ according to

$$k_{u|w}(t) = \frac{\sigma^2_{u|w}(t)}{b_{u|w}(t)}, \quad R_{u|w}(t) = \mathcal{R}\left(k_{u|w}(t)\right), \quad \mathcal{R}(k) = \frac{k}{2-k}. \tag{8}$$

As shown in Bermin and Holm (2021b) any bankruptcy-avoiding trading strategy holding the relative leverage risk process $k_{u|w}$ constant over time, at say a level $k \in (0, 2)$, has a maximal drawdown distribution given by the simple analytical formula

$$\mathbb{P}\left(\inf_{0\le t<\infty} \log \frac{X_{u|w}(t)}{X_{u|w}(0)} \le -n\mathcal{R}(k)\right) = e^{-n}, \quad n \ge 0. \tag{9}$$

While the formula provides an intuitive interpretation for the relative drawdown risk we do not necessarily require that the relative leverage risk process is kept constant over time. Instead we directly associate drawdown risk with the process $R_{u|w}$ and recall from Bermin and Holm (2021b) that this process shares many of the properties seen in coherent and convex risk measures, see Artzner et al. (1999), Föllmer and Schied (2002) for further details.

What makes the proposed framework compelling is that all quantities can be computed from the instantaneous covariance and rate of excess return processes. The dependency on the reference strategy can further be removed as explained below. Define the covariance matrix process V_0 of the investable assets, relative to the market numéraire, by the components

$$V_{0|n,m}(t) = V_{0|e_n,e_m}(t) = \frac{d}{dt}[\log P_{0|n}, \log P_{0|m}](t), \quad n, m \in \{1, \ldots, N\}. \tag{10}$$

Assume that the matrix V_0 is a.s. positive definite such that it generates an inner product of the form $\langle v, w\rangle_{V_0} = v'V_0 w$. It now follows, from Eqs. (1) and (3), that

$$V_{0|v,w}(t) = \langle v(t), w(t)\rangle_{V_0(t)}, \quad b_{0|w}(t) = \langle w_*(t), w(t)\rangle_{V_0(t)}, \quad w_*(t) = V_0^{-1}(t)b_0(t). \tag{11}$$

The particular trading strategy w_* is commonly known as the growth optimal Kelly strategy and it is easily seen, using Eqs. (6) and (11), that it can be characterized as

$$w_*(t) = \arg\max_{w(t)} \mu_{0|w}(t) = \arg\max_{w(t)} \mu_{u|w}(t), \quad \forall u. \tag{12}$$

The observation that the growth optimal Kelly strategy is independent of the reference strategy follows from the alternative representation $X_{u|w} = X_{0|w}/X_{0|u}$, which implies that $\mu_{u|w} = \mu_{0|w} - \mu_{0|u}$. Finally, we extend Eq. (11) to an arbitrary reference strategy as described below

Proposition 2.1. *For every reference strategy u, the instantaneous covariance process V_u and the rate of excess return process b_u equal*

$$V_{u|v,w}(t) = V_{0|v-u,w-u}(t), \quad b_{u|w}(t) = V_{0|w_*-u,w-u}(t).$$

Proof. The proof follows from straightforward calculations, see Bermin and Holm (2021b). □

This result shows that the minimal representation of the framework is given by the quantities (w_*, V_0). Once these quantities are specified everything else is computable. Having established the connection between an arbitrary reference strategy and the market numéraire we may now address topics such as: how can an investor increase the generalized Sharpe ratio of a portfolio. In order to answer such a question we follow Nielsen and Vassalou (2004) and apply a Taylor expansion to the term $s_{u|w+\varepsilon(v-u)}$. By the use of Lemma 2.1 we compute

$$s_{u|w+\varepsilon(v-u)}(t) = s_{u|w}(t) + \frac{\alpha_{u|v,w}(t)}{\sigma_{u|w}(t)}\varepsilon + \mathcal{O}(\varepsilon^2), \tag{13}$$

where

$$\alpha_{u|v,w}(t) = b_{u|v}(t) - \beta_{u|v,w}(t) b_{u|w}(t), \quad \beta_{u|v,w}(t) = \frac{V_{u|v,w}(t)}{V_{u|w}(t)} = \frac{\sigma_{u|v}(t)\rho_{u|v,w}(t)}{\sigma_{u|w}(t)}. \tag{14}$$

We recognise $\alpha_{u|v,w}$ as a generalized Jensen's alpha, see Jensen (1968), describing how the instantaneous excess return of the trading strategy v is risk adjusted with respect to the trading strategy w and the reference strategy u. The adjustment equals the product between the generalized risk parameter beta and the instantaneous excess return of the trading strategy w. It is apparent from Eq. (13) that for ε sufficiently small we can always improve the Sharpe ratio if Jensen's alpha is different from zero. While theoretically interesting this observation has, of course, limited practical applicability since only infinitesimal contributions are considered.

Remark 1. Note that we recover the standard definitions of alpha and beta when the reference strategy $u = \mathbf{0}$, and the numéraire asset can be identified with a locally risk-free bank account, expressed in terms of some interest rate process r.

In this paper we show how to calculate the optimal instantaneous Sharpe ratio when the opportunity set is enlarged. In doing so we use elements from Kelly portfolio theory and emphasize on the non-trivial geometry governing risk adjustments of returns. We stress that this is a static analysis, carried out for a fixed point in time, and consequently we often suppress the time dimension to facilitate the reading. For simplicity, we also focus on the case where the reference strategy corresponds to the market numéraire, i.e. $u = \mathbf{0}$, and provide details at a later stage on how to generalize the results derived.

3 Basic Geometry

In this section we give a very brief introduction to geometry, including tensors and tensor notation. For additional details see for instance Dodson and Poston (1991). The reason for choosing this path is that we sometimes need to study the geometry from the viewpoint of different coordinate systems. Consequently, it is beneficial to work with a coordinate free representation. The tensor notation further offers superior understanding, compared to the linear algebra matrix notation, in describing how the components transform with respect to linear transformations of the basis vectors. In order to easily distinguish components from basis vectors (and tensors) we write the latter ones in bold.

Throughout this paper let U be an N-dimensional vector space over $\mathbb{R}$ such that U is isomorphic to $\mathbb{R}^N$. A typical element of U is denoted by $\mathbf{w}$ and corresponds to a trading strategy at a given point in time. Expressed in terms of the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_N\}$ this means that $\mathbf{w} = w^1(t)\mathbf{e}_1 + \cdots + w^N(t)\mathbf{e}_N$ for some vector of components $w(t)$. Similarly, we let $\mathbf{w}_*$ denote the growth optimal Kelly vector with components $w_*(t)$ in the standard basis. We also fix the inner product $\mathbf{V}_0(\mathbf{v}, \mathbf{w}) = \langle v(t), w(t) \rangle_{V_0(t)}$, representing the random variable $V_{0|v,w}(t)$, and note that $\mathcal{H} = (U, \mathbf{V}_0)$ is a Hilbert space.

We further let U^* denote the dual vector space containing all linear forms on U. The elements of the dual space are referred to as covectors, or 1-forms, and a typical example is the instantaneous rate of excess return. We define the covector $\mathbf{b}_0(\mathbf{w}) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{w})$ such that it represent the random variable $b_{0|w}(t)$. Hence, $\mathbf{b}_0(\mathbf{e}_n)$ corresponds to the n'th term of the component vector $b_0(t) = (b_{0|1}(t), \ldots, b_{0|N}(t))'$ for the investable assets. The notion of the dual space is important throughout this work and from linear algebra we know that the dual space U^* is itself a vector space of the same dimension as U. Moreover, as U is finite dimensional the map into its double dual space U^{**} is a natural isomorphism; whence U^{**} can be identified with the original vector space. This means that we can also regard $\mathbf{w}$ as the linear form $\mathbf{w}(\mathbf{b}_0) = \mathbf{b}_0(\mathbf{w})$.

For any vector basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ of U, there exists a dual basis $\{\mathbf{u}^1, \ldots, \mathbf{u}^N\}$ for which $U^* = \mathrm{span}(\mathbf{u}^1, \ldots, \mathbf{u}^N)$. We express the canonical dual basis, using the Kronecker delta, according to

$$\mathbf{u}^i(\mathbf{u}_j) = \delta^i_j. \tag{15}$$

One notes that in the special case where the vector basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_N\}$ is orthonormal the canonical dual basis takes the same form as the vector basis, which consequently allows for considerable simplifications. In our situation, however, this is not the case. The standard basis is related to the investable assets, which are assumed to be correlated with each other. In order to further explain the relationship we apply a linear transformation to the standard basis. It follows, using Einstein summation for repeated indices, that if $\bar{\mathbf{e}}_i = A_i{}^j \mathbf{e}_j$ then $\bar{\mathbf{e}}^i = (A^{-1})_j{}^i \mathbf{e}^j$. We verify this statement using the linearity of covectors

$$\bar{\mathbf{e}}^i(\bar{\mathbf{e}}_j) = (A^{-1})_k{}^i \mathbf{e}^k (A_j{}^l \mathbf{e}_l) = A_j{}^l (A^{-1})_k{}^i \delta^k_l = A_j{}^l (A^{-1})_l{}^i = \delta^i_j, \tag{16}$$

Similarly, we notice that vector components also transform inversely to the coordinates as $\mathbf{w} = w^i(t)\mathbf{e}_i = \bar{w}^i(t)\bar{\mathbf{e}}_i = \bar{w}^i(t) A_i{}^j \mathbf{e}_j$ implies that $\bar{w}^i(t) = (A^{-1})_j{}^i w^j(t)$. For a covector, though, the components transform similar to the vector basis, i.e. with $\mathbf{b}_0 = b_{0|i}(t)\mathbf{e}^i = \bar{b}_{0|i}(t)\bar{\mathbf{e}}^i$ we obtain $\bar{b}_{0|i}(t) = A_i{}^j b_{0|j}(t)$. Furthermore, the asset-asset covariance matrix $V_0(t)$ generating the inner product $\mathbf{V}_0 = V_{0|i,j}(t)\mathbf{e}^i \otimes \mathbf{e}^j = \bar{V}_{0|i,j}(t)\bar{\mathbf{e}}^i \otimes \bar{\mathbf{e}}^j$ transforms according to $\bar{V}_{0|i,j}(t) = A_i{}^k A_j{}^l V_{0|k,l}(t)$.

The framework briefly outlined above is that of tensor analysis. The takeaway is that a tensor is always independent of the chosen basis but that the components change in such a way as to reflect the basis used. More formally, we regard a (p, q)-tensor $\mathbf{T}$ as an element of the space

$$\underbrace{U \otimes \cdots \otimes U}_{p} \otimes \underbrace{U^* \otimes \cdots \otimes U^*}_{q}, \tag{17}$$

such that $\mathbf{T}$ maps p covectors (recall that we identify U^{**} with U) and q vectors to $\mathbb{R}$ in a coordinate free and multilinear way. It is important to understand, however, that in order to compute the function value in $\mathbb{R}$ we must always choose a particular basis and identify the corresponding components.

$\mathbf{T}$	$\mathbf{T} \in$	(p, q)
$\mathbf{b}_0$	U^*	(0,1)
$\mathbf{w}$	U	(1,0)
$\mathbf{V}_0$	$U^* \otimes U^*$	(0,2)
$\mathbf{V}_0^{-1}$	$U \otimes U$	(2,0)
$\mathbf{P}_0$	$U \otimes U^*$	(1,1)

Table 1: Summary of main tensors. Note that other financial quantities, such as the relative leverage risk $\mathbf{k}_0(\mathbf{w}) = \mathbf{V}_0(\mathbf{w}, \mathbf{w})/\mathbf{b}_0(\mathbf{w})$, are typically not tensors due to the lack of multilinearity. For instance, although $\mathbf{k}_0(\lambda\mathbf{w}) = \lambda\mathbf{k}_0(\mathbf{w})$, we have $\mathbf{k}_0(\mathbf{v}+\mathbf{w}) \neq \mathbf{k}_0(\mathbf{v})+\mathbf{k}_0(\mathbf{w})$.

In Table 1 we highlight the main tensors used in this paper. With the abstract tensor notation we observe that the instantaneous rate of excess return covector $\mathbf{b}_0 = \mathbf{V}_0(\mathbf{w}_*)$ is the metric dual of the growth optimal Kelly vector. In other words, $\mathbf{w}_* \in \mathcal{H} = (U, \mathbf{V}_0)$ is the Riesz representation of $\mathbf{b}_0 \in \mathcal{H}^* = (U^*, \mathbf{V}_0^{-1})$, where the inner product $\mathbf{V}_0^{-1} = (V_0^{-1}(t))^{ij}\mathbf{e}_i \otimes \mathbf{e}_j$ is generated by the inverse covariance matrix $V_0^{-1}(t)$ of the investable assets, such that $\|\mathbf{w}_*\|_{\mathcal{H}} = \|\mathbf{b}_0\|_{\mathcal{H}^*}$. Similarly, we can also write $\mathbf{w}_* = \mathbf{V}_0^{-1}(\mathbf{b}_0)$, with the interpretation that $\mathbf{w}_*$ is an element of the double dual space $U^{**} \cong U$, such that $\mathbf{w}_*(\mathbf{e}^n) = \mathbf{e}^n(\mathbf{w}_*)$ equals the n'th component of $w_*(t)$. The $(1,1)$-tensor $\mathbf{P}_0 = P^i_{0|j}(t)\mathbf{e}_i \otimes \mathbf{e}^j$ is a projection operator mapping a covector and a vector to $\mathbb{R}$. More commonly, though, we regard it as a map from either U onto U or from U^* onto U^*. Finally, let us mention that we have chosen to represent other financial key quantities with the same notation, although they are not tensors. For instance, we let $\mathbf{s}_0(\mathbf{w}) = \mathbf{b}_0(\mathbf{w})/\sqrt{\mathbf{V}_0(\mathbf{w}, \mathbf{w})}$ denote the instantaneous Sharpe ratio and note that this quantity is truly speaking not a tensor due to the non-linear scaling $\mathbf{s}_0(\lambda\mathbf{w}) = \mathrm{sign}(\lambda)\mathbf{s}_0(\mathbf{w})$.

For the remainder of this section we present additional results related to the growth optimal Kelly vector, with the purpose both to motivate the use of tensors and to present a framework suitable for geometric analysis.

3.1 Absence of Arbitrage

In order to highlight the power of tensor analysis we provide an enlightening example of when it is important to consider vectors rather than simply components for a particular basis.

Recall that in Section 2 we presented the portfolio theory directly in terms of the components corresponding to the standard basis. However, we did not explicitly specify the source of randomness driving the evolution of the investable assets. By assuming that these assets are $\mathbb{F}$-adapted continuous processes we may write

$$\frac{dP_{0|n}(t)}{P_{0|n}(t)} = b_{0|n}(t)dt + \Sigma_{0|n,m}(t)dW^m(t), \quad n \in \{1, \ldots, N\}, \tag{18}$$

for some standard Brownian motion $W = (W^1, \ldots, W^M)'$ and some $\mathbb{F}$-adapted, $\mathbb{R}^{N\times M}$-valued, volatility process Σ_0. The components of the investable asset-asset covariance matrix $V_0(t)$ then equals $\Sigma_0(t)\Sigma_0'(t)$. It is well known that absence of arbitrage implies the existence of an $\mathbb{F}$-adapted process $\theta = (\theta^1, \ldots, \theta^M)'$, see for instance Karatzas and Shreve (1999), such that

$$\Sigma_{0|n,m}(t)\theta^m(t) = b_{0|n}(t). \tag{19}$$

We call θ the market price of risk process and notice that in a complete market, where $M = N$, this process equals $\theta(t) = \Sigma_0^{-1}(t)b_0(t)$. Consequently, in a complete market it follows, from Eq. (11), that we can express the growth optimal Kelly strategy as

$$\mathbf{w}_* = w_*^i(t)\mathbf{e}_i = \theta^a(t)(\Sigma_0^{-1}(t))_a{}^j\mathbf{e}_j = \theta^a(t)\bar{\mathbf{e}}_a. \tag{20}$$

Moreover, since $\{\bar{\mathbf{e}}_1, \ldots, \bar{\mathbf{e}}_N\}$ is a basis of U we see that $\theta^a(t)\bar{\mathbf{e}}_a$ naturally describes the market price of risk vector $\mathbf{\Theta} \in U$. We summarize the observations below.

Theorem 3.1. *In a complete market, where $M = N$, the market price of risk vector $\mathbf{\Theta}$ is identical to the growth optimal Kelly vector $\mathbf{w}_*$. That is, with $\mathbf{\Theta} = \theta^a(t)\bar{\mathbf{e}}_a$, the components* *and the basis vectors relate according to*

$$\theta^a(t) = w_*^j(t)(\Sigma_0(t))_j{}^a, \qquad \bar{\mathbf{e}}_a = (\Sigma_0^{-1}(t))_a{}^j \mathbf{e}_j,$$
$$w_*^i(t) = \theta^a(t)(\Sigma_0^{-1}(t))_a{}^i, \qquad \mathbf{e}_i = (\Sigma_0(t))_i{}^a \bar{\mathbf{e}}_a,$$

such that

$$\mathbf{s}_0(\mathbf{w}_*) = \|\mathbf{w}_*\|_{\mathcal{H}} = \|\boldsymbol{\Theta}\|_{\mathcal{H}} = \|\theta(t)\|_{\mathbb{R}^N}.$$

Note further that we can always choose a new orthonormal basis $\{\check{\mathbf{e}}_1, \ldots, \check{\mathbf{e}}_N\}$*, through a standard orthogonal coordinate transformation, such that*

$$\boldsymbol{\Theta} = \mathbf{s}_0(\mathbf{w}_*)\check{\mathbf{e}}_1, \quad \check{\mathbf{e}}_1 = \frac{\mathbf{w}_*}{\|\mathbf{w}_*\|_{\mathcal{H}}}.$$

Proof. Given that $\mathbf{w}_* = \boldsymbol{\Theta}$ we need to show that $\|\mathbf{w}_*\|_{\mathcal{H}} = \mathbf{s}_0(\mathbf{w}_*)$ and $\|\boldsymbol{\Theta}\|_{\mathcal{H}} = \|\theta(t)\|_{\mathbb{R}^N}$. Direct calculations using Eqs. (7) and (11) yield

$$\mathbf{s}_0^2(\mathbf{w}_*) = \frac{\mathbf{b}_0^2(\mathbf{w}_*)}{\mathbf{V}_0(\mathbf{w}_*, \mathbf{w}_*)} = \mathbf{V}_0(\mathbf{w}_*, \mathbf{w}_*) = \|\mathbf{w}_*\|_{\mathcal{H}}^2.$$

Furthermore, since the components of $\mathbf{w}$, with respect to the standard basis, satisfy $w_*'(t) = \theta'(t)\Sigma_0^{-1}(t)$ and $V_0(t)$ admits the decomposition $\Sigma_0(t)\Sigma_0'(t)$ it follows that

$$\|\mathbf{w}_*\|_{\mathcal{H}}^2 = \langle w_*(t), w_*(t)\rangle_{V_0(t)} = w_*'(t)V_0(t)w_*(t) = \theta'(t)\theta(t),$$

which we recognize as the square of the Euclidean norm in $\mathbb{R}^N$. □

By taking a geometric approach we identify the growth optimal Kelly vector with the market price of risk vector in a complete market. The key observation is that in algebra and analysis the latter vector is typically expressed using components from a basis different from the standard basis, which muddies the water and hides the fact that the length of the vector equals its instantaneous Sharpe ratio. With this introduction to Kelly trading we proceed by investigating how to characterize the growth optimal Kelly vector on subspaces.

3.2 The Opportunity Set and Projections

So far we have considered the opportunity set to consist of N numéraire based investable assets. The first observation to be made is that this dimension is local in time since new assets might be available for investment in the future while other assets might cease to exist for various reasons. However, for a given point in time, the dimension of the opportunity set can also vary from investor to investor and below we aim to clarify the geometry governing such reductions or expansions.

The approach we follow is to consider a subspace $U_1 \subseteq U$. Any vector $\mathbf{w} \in U_1$ can be expressed as $\mathbf{w} = w^j_{\mathbf{v}}(t)\mathbf{v}_j$ for a given basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_{N_1}\}$, $N_1 \leq N$, of U_1. Hence, the N_1 investable assets of the opportunity set U_1 are linear combinations of the N investable assets in U. We can further translate the representation to the standard basis of U, by setting $\mathbf{v}_j = v^i_j \mathbf{e}_i$, such that $\mathbf{w} = w^i(t)\mathbf{e}_i$ with $w^i(t) = w^j_{\mathbf{v}}(t)v^i_j$. Below we show how to characterize the growth optimal Kelly vector on U_1, defined by

$$\mathbf{w}_*[U_1] = \arg\max_{\mathbf{w} \in U_1} \mu_0(\mathbf{w}), \tag{21}$$

in terms of $\mathbf{w}_*$. However, first we introduce a technical results.

Lemma 3.2. *Given a subspace $\mathcal{H}_1 \subseteq \mathcal{H}$. The orthogonal projection of a vector $\mathbf{w} \in \mathcal{H}$ onto $\mathcal{H}_1 = (U_1, \mathbf{V}_0)$ is unique and satisfies*

$$\mathbf{V}_0(\mathbf{w}, \mathbf{P}_{0|U_1}(\mathbf{x})) = \mathbf{V}_0(\mathbf{P}_{0|U_1}(\mathbf{w}), \mathbf{x}) = \mathbf{V}_0(\mathbf{P}_{0|U_1}(\mathbf{w}), \mathbf{P}_{0|U_1}(\mathbf{x})), \quad \forall \mathbf{x} \in \mathcal{H}.$$

Furthermore, the orthogonal projection admits the representation

$$\mathbf{P}_{0|U_1}(\mathbf{w}) = \sum_{i \geq 1} \frac{\mathbf{V}_0(\mathbf{w}, \mathbf{v}_i)}{\mathbf{V}_0(\mathbf{v}_i, \mathbf{v}_i)} \mathbf{v}_i,$$

for any orthogonal sequence $\{\mathbf{v}_i\}_{i \geq 1}$ spanning U_1.

Proof. For details about the proof we refer to Luenberger (1997). □

It is worth mentioning that the functional representation of the orthogonal projection is more complicated when expanded in a non-orthogonal basis; a topic we return to later in this paper. With that being said, we now return to the growth optimal Kelly vector and highlight the financial connection.

Theorem 3.3. *For $\mathcal{H}_1 = (U_1, \mathbf{V}_0)$ let $\mathcal{H}_1 \subseteq \mathcal{H}$. Then*

$$\mathbf{w}_*[U_1] = \mathbf{P}_{0|U_1}(\mathbf{w}_*), \quad \|\mathbf{w}_*[U_1]\|_{\mathcal{H}} = \mathbf{s}_0(\mathbf{w}_*[U_1]).$$

Proof. Let $\{\mathbf{v}_i\}_{i\leq N_1}$, $N_1 = \dim(U_1)$, be an orthogonal sequence spanning U_1 such that any vector $\mathbf{w}$ in U_1 takes the form $\mathbf{w} = \lambda^j \mathbf{v}_j$. It now follows from Eqs. (6) and (11) that

$$\mu_0(\lambda^j \mathbf{v}_j) = \lambda^j \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_j) - \frac{1}{2}\lambda^j \lambda^k \mathbf{V}_0(\mathbf{v}_j, \mathbf{v}_k).$$

Hence, the rate of excess logarithmic return is maximal when

$$0 = \frac{\partial}{\partial \lambda^i}\mu_0(\lambda^j \mathbf{v}_j) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_i) - \lambda^k \mathbf{V}_0(\mathbf{v}_i, \mathbf{v}_k).$$

Since $\{\mathbf{v}_i\}_{i\leq N_1}$ is an orthogonal sequence we see that $\lambda^i = \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_i)/\mathbf{V}_0(\mathbf{v}_i, \mathbf{v}_i)$. The first part of the proof follows by identifying the terms with those in Lemma 3.2. Having identified the growth optimal Kelly vector on a subspace as a projection we again apply Lemma 3.2 to obtain

$$\mathbf{b}_0(\mathbf{w}_*[U_1]) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{P}_{0|U_1}(\mathbf{w}_*)) = \mathbf{V}_0(\mathbf{P}_{0|U_1}(\mathbf{w}_*), \mathbf{P}_{0|U_1}(\mathbf{w}_*)) = \|\mathbf{w}_*[U_1]\|^2_{\mathcal{H}},$$

from which the proof concludes. □

A different explanation can be seen from the expression $\mu_0(\mathbf{v}) = \frac{1}{2}(\|\mathbf{w}_*\|^2_{\mathcal{H}} - \|\mathbf{w}_* - \mathbf{v}\|^2_{\mathcal{H}})$, which shows that the local maximum is attained at the point with minimal distance to the growth optimal Kelly vector. Hence, for any subspace, the line from this unique point to $\mathbf{w}_*$ is orthogonal to the subspace and therefore coincides with the orthogonal projection of $\mathbf{w}_*$ onto the subspace.

Remark 2. By setting $\mathbf{b}_0[U_1^*] = \mathbf{V}_0(\mathbf{w}_*[U_1])$ one notes, from Lemma 3.2, that

$$\mathbf{b}_0[U_1^*](\mathbf{v}) = \mathbf{V}_0(\mathbf{P}_{0|U_1}(\mathbf{w}_*), \mathbf{v}) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{P}_{0|U_1}(\mathbf{v})) = \mathbf{b}_0(\mathbf{v}), \quad \mathbf{v} \in \mathcal{H}_1 \subseteq \mathcal{H}. \tag{22}$$

The interpretation is that $\mathbf{b}_0[U_1^*]$ can be expressed in any dual basis spanning U_1^*, while $\mathbf{b}_0$

must be expanded in any dual basis spanning U^*. Similarly, we sometimes write $\mathbf{V}_0[U_1^*]$ when to emphasize that the inner product can be expanded using a dual basis spanning U_1^*. While the components of the expansions change for every chosen basis, it is important to remember that the mapping to the real numbers do not. Consequently, we often omit the notion of subspace for ease of readability.

We further see that we can regard the Hilbert space $\mathcal{H}$, corresponding to the investable assets for a given investor, as a subspace of the Hilbert space $\bar{\mathcal{H}} = (\bar{U}, \mathbf{V}_0)$ representing all the world's assets. What this means is that when analyzing optimal portfolio allocations, for a particular investor, we only have to consider the covariance structure of the investable assets for that investor. This follows as, restricted to a subspace U_1, we only need to find the components of $\mathbf{V}_0$ for a given basis (meaning the investable assets) of U_1. It is quite remarkable that we can equate the growth optimal Kelly vector on any subspace with a projection of the worldwide growth optimal Kelly vector. This feature further implies that a growth optimal Kelly vector can be expressed as a nested sequence of projections

$$\mathbf{w}_*[U_K] = \mathbf{P}_{0|U_K} \cdots \mathbf{P}_{0|U_1}(\mathbf{w}_*), \quad U_K \subset \cdots \subset U_1. \tag{23}$$

The financial interpretation of such a nested sequence can best be appreciated from a simple application of Cauchy-Schwarz inequality; stating that $\|\mathbf{P}_{0|U_k}(\mathbf{w})\|_{\mathcal{H}} \leq \|\mathbf{w}\|_{\mathcal{H}}$ for all $\mathbf{w} \in \mathcal{H}$. Consequently $\|\mathbf{w}_*[U_k]\|_{\mathcal{H}} \leq \|\mathbf{w}_*[U_{k-1}]\|_{\mathcal{H}}$, which implies (see Theorem 3.3) that $\mathbf{s}_0(\mathbf{w}_*[U_k]) \leq \mathbf{s}_0(\mathbf{w}_*[U_{k-1}])$. Hence, at each time we reduce the dimension of the investable assets, for instance by replacing some assets by a mutual fund, the maximal instantaneous Sharpe ratio is reduced.

3.3 Level Sets, Correlations and Reflections

Modern portfolio theory is largely based on the geometric principle that the level sets of the instantaneous Sharpe ratio are cones. In other words, if we set $C_s = \{\mathbf{w} \in \mathbb{R}^N : \mathbf{s}_0(\mathbf{w}) = s\}$, then for each $\mathbf{w} \in C_s$, and positive scalar $\lambda > 0$, we have $\lambda\mathbf{w} \in C_s$. Markowitz (1952) and Tobin (1958) used this property to derive the well-known efficient mean-variance frontier; characterized by the set of trading strategies for which the Sharpe ratio is maximal. Kelly (1956) and Latané (1959), however, argued that by leveraging too hard (that is using a too high λ) the logarithmic excess return, as opposed to the excess return, eventually becomes negative. Following, Bermin and Holm (2021b) we illustrate this feature using the concept of relative leverage/drawdown risk

$$\mathbf{b}_0(\mathbf{w}) = \frac{1}{\mathbf{k}_0(\mathbf{w})}\sigma_0^2(\mathbf{w}), \quad \mu_0(\mathbf{w}) = \left(\frac{1}{\mathbf{k}_0(\mathbf{w})} - \frac{1}{2}\right)\sigma_0^2(\mathbf{w}) = \frac{1}{2\mathcal{R}(\mathbf{k}_0(\mathbf{w}))}\sigma_0^2(\mathbf{w}). \tag{24}$$

Hence, the instantaneous excess return is strictly positive if and only if $\mathbf{k}_0(\mathbf{w}) > 0$, while the instantaneous logarithmic excess return is strictly positive if and only if $0 < \mathbf{k}_0(\mathbf{w}) < 2$.

In order to visualize the framework geometrically we first claim that, for $\mathbf{w} \in \mathcal{H}$, the level sets of $\sigma_0(\mathbf{w})$, $\mu_0(\mathbf{w})$ and $\mathbf{k}_0(\mathbf{w})$ are spheres of dimension $N-1$, while the level sets of $\mathbf{b}_0(\mathbf{w})$ are hyperplanes of dimension $N-1$. Furthermore, there exists a sphere of dimension $N-2$, for which all trading strategies are equivalent with respect to the quantities just mentioned.

Proposition 3.4. *The various level sets can be characterized by*

Level	*Topology*	*Center*	*Radius*
$\mathbf{b}_0(\mathbf{w}) = b$	$\mathbb{R}^{N-1}$	$\frac{b}{\mathbf{s}_*^2}\mathbf{w}_*$	$-$
$\sigma_0(\mathbf{w}) = \sigma$	S^{N-1}	0	σ
$\mu_0(\mathbf{w}) = \mu$	S^{N-1}	$\mathbf{w}_*$	$\mathbf{s}_*\sqrt{1 - \frac{2\mu}{\mathbf{s}_*^2}}$
$\mathbf{k}_0(\mathbf{w}) = k$	S^{N-1}	$\frac{1}{2}k\mathbf{w}_*$	$\frac{1}{2}k\mathbf{s}_*$
$\mathbf{b}_0(\mathbf{w}) = b$, $\sigma_0(\mathbf{w}) = \sigma$	S^{N-2}	$\frac{b}{\mathbf{s}_*^2}\mathbf{w}_*$	$\sigma\sqrt{1 - \left(\frac{b}{\sigma\mathbf{s}_*}\right)^2}$

where we have set $\mathbf{s}_* = \mathbf{s}_0(\mathbf{w}_*)$ *for convenience. Note further that the joint levels sets of* $(\mathbf{b}_0, \sigma_0)$ *imply level sets for* $(\mu_0, \mathbf{k}_0)$.

Proof. We, unconventionally, express the quantities using the norm on $\mathcal{H}$ according to

$$\sigma_0^2(\mathbf{w}) = \|\mathbf{w}\|_{\mathcal{H}}^2, \quad \mu_0(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}_*\|_{\mathcal{H}}^2 - \frac{1}{2}\|\mathbf{w} - \mathbf{w}_*\|_{\mathcal{H}}^2,$$
$$\mathbf{k}_0^2(\mathbf{w}) = \frac{4}{\|\mathbf{w}_*\|_{\mathcal{H}}^2}\|\mathbf{w} - \frac{1}{2}\mathbf{k}_0(\mathbf{w})\mathbf{w}_*\|_{\mathcal{H}}^2.$$

We also note from Lemma 3.2 that

$$\mathbf{w}_{\|} = \mathbf{P}_{0|\,\mathrm{span}(\mathbf{w}_*)}(\mathbf{w}) = \frac{\mathbf{V}_0(\mathbf{w}, \mathbf{w}_*)}{\mathbf{V}_0(\mathbf{w}_*, \mathbf{w}_*)}\mathbf{w}_* = \frac{\mathbf{b}_0(\mathbf{w})}{\|\mathbf{w}_*\|^2_{\mathcal{H}}}\mathbf{w}_*,$$

such that, with $\mathbf{w}_\perp = \mathbf{w} - \mathbf{w}_{\|}$, we have

$$\|\mathbf{w}_\perp\|^2_{\mathcal{H}} = \|\mathbf{w} - \mathbf{w}_{\|}\|^2_{\mathcal{H}} = \sigma_0^2(\mathbf{w}) - \frac{\mathbf{b}_0^2(\mathbf{w})}{\|\mathbf{w}_*\|^2_{\mathcal{H}}}.$$

Finally, we identify the center point and the radius of the expressions. We also recall that the norm of the growth optimal Kelly vector equals its instantaneous Sharpe ratio as shown in Theorem 3.3. The proof concludes from the observation that $(\mu_0, \mathbf{k}_0)$ can be expressed in terms of $(\mathbf{b}_0, \sigma_0)$. □

The importance of the growth optimal Kelly vector can be explained from the observation that the instantaneous excess return is invariant with respect to trading strategies orthogonal to $\mathbf{w}_*$. That is, with $\mathbf{v} = \mathbf{v}_{\|} + \mathbf{v}_\perp$, where $\mathbf{v}_{\|}$ and $\mathbf{w}_*$ are collinear while $\mathbf{v}_\perp$ and $\mathbf{w}_*$ are perpendicular, one sees that

$$\mathbf{b}_0(\mathbf{v}) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{\|} + \mathbf{v}_\perp) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{\|}) + \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_\perp) = \mathbf{b}_0(\mathbf{v}_{\|}). \tag{25}$$

However, since the volatility increases with $\mathbf{v}_\perp$, through the formula $\sigma_0^2(\mathbf{v}) = \sigma_0^2(\mathbf{v}_{\|}) + \sigma_0^2(\mathbf{v}_\perp)$, it is clear that the instantaneous Sharpe ratio is maximal for trading strategies collinear to $\mathbf{w}_*$. Hence, as can be seen from Fig. 1, the instantaneous efficient mean-variance frontier (minimal variance for a fixed excess return) consists of all vectors collinear to the growth optimal Kelly vector. These vectors can, however, equally be represented by different constraint optimization problems, such as minimal relative leverage risk for a fixed (logarithmic) excess return, to give an example.

We proceed by considering the projection of the growth optimal Kelly vector on the subspace spanned by a single vector $\mathbf{v}$. By the use of Eq. (8) and Lemma 3.2 we define

$$\hat{\mathbf{v}} = \frac{1}{\mathbf{k}_0(\mathbf{v})}\mathbf{v} = \frac{\mathbf{V}_0(\mathbf{w}_*, \mathbf{v})}{\mathbf{V}_0(\mathbf{v}, \mathbf{v})}\mathbf{v} = \mathbf{P}_{0|\,\mathrm{span}(\mathbf{v})}(\mathbf{w}_*) = \mathbf{w}_*[\mathrm{span}(\mathbf{v})]. \tag{26}$$

Bermin and Holm (2021b) call trading strategies generated in this way for generalized

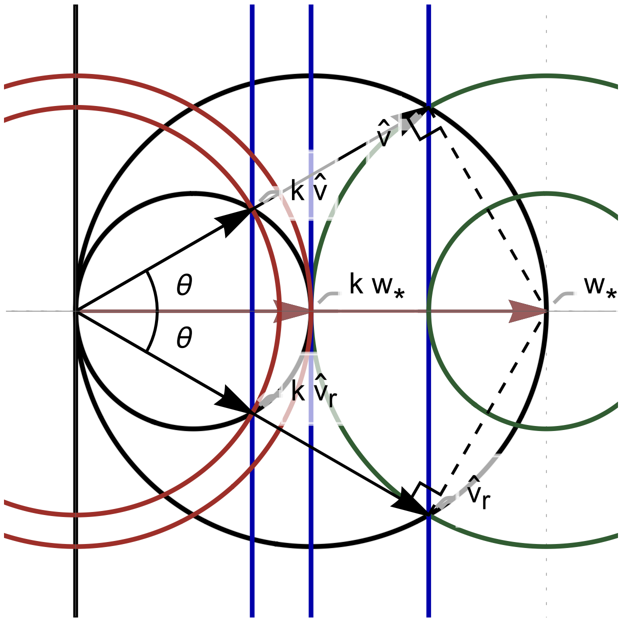


Figure 1: This figure shows the vectors $\mathbf{v} = k\hat{\mathbf{v}}$ and $\mathbf{v}_r = k\hat{\mathbf{v}}_r$, where the latter is a reflection of the first through the line spanned by the growth optimal Kelly vector $\mathbf{w}_*$. We also highlight the level sets of $\mathbf{k}_0$ (black), $\mathbf{b}_0$ (blue), μ_0 (green) and those of σ_0 (red).

growth optimal Kelly strategies and show that these strategies have the same relative drawdown/leverage risk as the growth optimal Kelly strategy. Their proof is a direct consequence of the simple relationships $\mathbf{k}_0(\lambda\mathbf{v}) = \lambda\mathbf{k}_0(\mathbf{v})$ and $\mathbf{k}_0(\mathbf{w}_*) = 1$. Since $\hat{\mathbf{v}}$ is the orthogonal projection of the growth optimal Kelly vector onto $U_1 = \mathrm{span}(\mathbf{v})$ the vector $\mathbf{w}_* - \hat{\mathbf{v}}$ is further perpendicular to $\hat{\mathbf{v}}$. Consequently, as shown in Fig 1, the angle between the vectors $\mathbf{w}_*$ and $\hat{\mathbf{v}}$ satisfy $\cos\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}} = \|\hat{\mathbf{v}}\|_{\mathcal{H}}/\|\mathbf{w}_*\|_{\mathcal{H}}$, The financial interpretation of the angle between vectors is the correlation and through the relationship $\rho_0(\mathbf{v},\mathbf{w}) = \cos\varphi_{\mathbf{v},\mathbf{w}}$ we obtain the following result.

Theorem 3.5. *For $\mathcal{H}_1 = (U_1, \mathbf{V}_0)$ let $\mathbf{v} \in \mathcal{H}_1 \subseteq \mathcal{H}$. Then*

$$\rho_0(\mathbf{v}, \mathbf{w}_*[U_1]) = \frac{\mathbf{s}_0(\mathbf{v})}{\mathbf{s}_0(\mathbf{w}_*[U_1])}.$$

Proof. Straightforward calculations, using Lemma 3.2, yield

$$\rho_0(\mathbf{v}, \mathbf{w}_*[U_1]) = \frac{\mathbf{V}_0(\mathbf{v}, \mathbf{P}_{0|U_1}(\mathbf{w}_*))}{\|\mathbf{v}\|_{\mathcal{H}}\|\mathbf{w}_*[U_1]\|_{\mathcal{H}}} = \frac{\mathbf{V}_0(\mathbf{P}_{0|U_1}(\mathbf{v}), \mathbf{w}_*)}{\|\mathbf{v}\|_{\mathcal{H}}\|\mathbf{w}_*[U_1]\|_{\mathcal{H}}} = \frac{\mathbf{b}_0(\mathbf{P}_{0|U_1}(\mathbf{v}))}{\sigma_0(\mathbf{v})\|\mathbf{w}_*[U_1]\|_{\mathcal{H}}}.$$

Hence, for $\mathbf{v} \in \mathcal{H}_1 \subseteq \mathcal{H}$, the proof concludes by the use of Theorem 3.3. $\square$

Since the correlation between any vector and the growth optimal Kelly vector equals the ratio of their instantaneous Sharpe ratios, we again see that $|\mathbf{s}_0(\mathbf{v})| \leq \mathbf{s}_0(\mathbf{w}_*[U_1])$ for any vector $\mathbf{v} \in U_1$. Furthermore, the correlation is in fact bounded by the various Sharpe ratios as shown below.

Corollary 3.6. *For $\mathcal{H}_1 = (U_1, \mathbf{V}_0)$ let $\mathbf{v}, \mathbf{w} \in \mathcal{H}_1 \subseteq \mathcal{H}$. Then*

$$\left|\rho_0(\mathbf{v},\mathbf{w}) - \frac{\mathbf{s}_0(\mathbf{v})\mathbf{s}_0(\mathbf{w})}{\mathbf{s}_0^2(\mathbf{w}_*[U_1])}\right| \leq \sqrt{\left(1 - \frac{\mathbf{s}_0^2(\mathbf{v})}{\mathbf{s}_0^2(\mathbf{w}_*[U_1])}\right)\left(1 - \frac{\mathbf{s}_0^2(\mathbf{w})}{\mathbf{s}_0^2(\mathbf{w}_*[U_1])}\right)},$$

with equality if $\dim(U_1) = 2$.

Proof. Define the vectors $\mathbf{v}_\perp = \mathbf{v} - \mathbf{P}_{0|\operatorname{span}(\mathbf{w}_*[U_1])}(\mathbf{v})$ and $\mathbf{w}_\perp = \mathbf{w} - \mathbf{P}_{0|\operatorname{span}(\mathbf{w}_*[U_1])}(\mathbf{w})$. Direct calculations, using Lemma 3.2, then show that

$$\mathbf{V}_0(\mathbf{v}_\perp, \mathbf{w}_\perp) = \sigma_0(\mathbf{v})\sigma_0(\mathbf{w})\left(\rho_0(\mathbf{v},\mathbf{w}) - \rho_0(\mathbf{w}_*[U_1],\mathbf{v})\rho_0(\mathbf{w}_*[U_1],\mathbf{w})\right),$$

which yields

$$\rho_0(\mathbf{v}_\perp, \mathbf{w}_\perp) = \frac{\rho_0(\mathbf{v},\mathbf{w}) - \rho_0(\mathbf{w}_*[U_1],\mathbf{v})\rho_0(\mathbf{w}_*[U_1],\mathbf{w})}{\sqrt{1-\rho_0^2(\mathbf{w}_*[U_1],\mathbf{v})}\sqrt{1-\rho_0^2(\mathbf{w}_*[U_1],\mathbf{w})}}.$$

Since $|\rho_0(\mathbf{v}_\perp, \mathbf{w}_\perp)| \leq 1$ the first part of the proof follows from Theorem 3.5.

We further note that if $\dim(U_1) = 2$ then $\mathbf{v}, \mathbf{w}$ and $\mathbf{w}_*[U_1]$ lie in the same plane. This means that the angle $\varphi_{\mathbf{v},\mathbf{w}} = \varphi_{\mathbf{w}_*[U_1],\mathbf{v}} + \varphi_{\mathbf{w}_*[U_1],\mathbf{w}}$, or $\varphi_{\mathbf{v},\mathbf{w}} = 2\pi - \varphi_{\mathbf{w}_*[U_1],\mathbf{v}} - \varphi_{\mathbf{w}_*[U_1],\mathbf{w}}$, or $\varphi_{\mathbf{v},\mathbf{w}} = \pm(\varphi_{\mathbf{w}_*[U_1],\mathbf{v}} - \varphi_{\mathbf{w}_*[U_1],\mathbf{w}})$, such that $\varphi_{\mathbf{v},\mathbf{w}} \in [0,\pi]$. By inspecting each case, we find that

$$\cos\varphi_{\mathbf{v},\mathbf{w}} = \cos\varphi_{\mathbf{w}_*[U_1],\mathbf{v}}\cos\varphi_{\mathbf{w}_*[U_1],\mathbf{w}} \mp \sin\varphi_{\mathbf{w}_*[U_1],\mathbf{v}}\sin\varphi_{\mathbf{w}_*[U_1],\mathbf{w}},$$

where the sign preceding the sine functions is negative for the first two representations of $\varphi_{\mathbf{v},\mathbf{w}}$ and positive for the latter two. The proof now follows from Theorem 3.5. $\square$

In Fig. 1 we also plot the reflection of the vector $\mathbf{v}$ with respect to the growth optimal Kelly vector. Hence, by setting $\mathbf{v}_r = \mathbf{v}_\| - \mathbf{v}_\perp$, such that

$$\mathbf{v}_r = 2\mathbf{v}_\| - \mathbf{v} = 2\mathbf{P}_{0|\operatorname{span}(\mathbf{w}_*)}(\mathbf{v}) - \mathbf{v} = 2\frac{\mathbf{V}_0(\mathbf{w}_*,\mathbf{v})}{\mathbf{V}_0(\mathbf{w}_*,\mathbf{w}_*)}\mathbf{w}_* - \mathbf{v}, \tag{27}$$

straightforward calculations yield

$$\mathbf{b}_0(\mathbf{v}_r) = \mathbf{b}_0(\mathbf{v}), \quad \mathbf{V}_0(\mathbf{v}_r, \mathbf{v}_r) = \mathbf{V}_0(\mathbf{v}, \mathbf{v}). \tag{28}$$

From these expressions it follows that also the instantaneous: excess logarithmic return, Sharpe ratio and relative drawdown/leverage risk are invariant quantities. The fact that we can, in general, identify two distinct trading strategies with identical local characteristics is a result of great importance in order to fully understand the widely used mean-variance framework. By construction we also note that

$$\mathbf{V}_0(\mathbf{v}_r, \mathbf{w}) = \mathbf{V}_0(\mathbf{w}_r, \mathbf{v}) = 2\frac{\mathbf{V}_0(\mathbf{w}_*, \mathbf{v})\mathbf{V}_0(\mathbf{w}_*, \mathbf{w})}{\mathbf{V}_0(\mathbf{w}_*, \mathbf{w}_*)} - \mathbf{V}_0(\mathbf{v}, \mathbf{w}), \tag{29}$$

which, together with Theorems 3.3 and 3.5, implies the identity

$$\rho_0(\mathbf{v}_r, \mathbf{w}) = \rho_0(\mathbf{w}_r, \mathbf{v}) = 2\frac{\mathbf{s}_0(\mathbf{v})\mathbf{s}_0(\mathbf{w})}{\mathbf{s}_0^2(\mathbf{w}_*)} - \rho_0(\mathbf{v}, \mathbf{w}). \tag{30}$$

Hence, for every pair of correlated trading strategies $(\mathbf{v}, \mathbf{w})$ we can always find new pairs $(\mathbf{v}_r, \mathbf{w})$ and $(\mathbf{v}, \mathbf{w}_r)$ with modified correlation but with otherwise identical characteristics. Note also that in the particular case where $\mathbf{w} = \mathbf{v}$, we obtain

$$\rho_0(\mathbf{v}_r, \mathbf{v}) = 2\rho_0^2(\mathbf{w}_*, \mathbf{v}) - 1 = 2\cos^2\varphi_{\mathbf{w}_*,\mathbf{v}} - 1 = \cos 2\varphi_{\mathbf{w}_*,\mathbf{v}}, \tag{31}$$

which confirms that the angle $\varphi_{\mathbf{v}_r,\mathbf{v}} = 2\varphi_{\mathbf{w}_*,\mathbf{v}}$ as illustrated in Fig. 1.

We have shown that the only trading strategies which are locally unique, in the sense mentioned above, are the so-called (fractional) Kelly strategies, $\mathbf{w} = k\mathbf{w}_*$, first introduced in MacLean, Ziemba and Blazenko (1992). For these trading strategies, characterized by having maximal instantaneous squared Sharpe ratio, one easily verifies that

$$\mu_0(k\mathbf{w}_*) = \frac{1}{2}k(2-k)\mathbf{s}_0^2(\mathbf{w}_*), \quad \sigma_0^2(k\mathbf{w}_*) = k^2\mathbf{s}_0^2(\mathbf{w}_*). \tag{32}$$

Consequently, a Kelly strategy is efficient if the relative leverage risk $\mathbf{k}_0(k\mathbf{w}_*) = k \in [0, 1]$, since otherwise we can always lower the volatility without reducing the logarithmic excess return. Finally, we briefly discuss how the geometric framework can further be used to visualize trade-offs between risk and return. For instance, from Fig. 1 we deduce how to lower the relative leverage/drawdown risk of an arbitrary trading strategy, at no expense on the logarithmic excess return, by employing an efficient Kelly strategy. We illustrate the approach by calculating the fraction $k \in [0,1]$ such that $\mu_0(k\mathbf{w}_*) = \mu_0(\hat{\mathbf{v}})$, using geometric principles only. One sees that the radius of the circle describing the level sets of the logarithmic excess return can be expressed in the two different ways: $\sin\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}}\|\mathbf{w}_*\|_{\mathcal{H}}$ and $(1-k)\|\mathbf{w}_*\|_{\mathcal{H}}$. Hence, with

$$k = 1 - \sin\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}} = 1 - \sqrt{1-\cos^2\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}}} = 1 - \sqrt{1-\rho_0^2(\mathbf{w}_*,\hat{\mathbf{v}})}, \tag{33}$$

the relative leverage/drawdown risk is reduced from $\mathbf{k}_0(\hat{\mathbf{v}}) = 1$ to $\mathbf{k}_0(k\mathbf{w}_*) = k \leq 1$, without affecting the excess logarithmic return. In much the same way it follows that, for a fixed logarithmic excess return, the trading strategies with lowest volatility are the Kelly strategies. This observation is a direct consequence of Proposition 3.4; stating that the level sets of the volatility are spheres centered at origo.

4 Risk Adjusted Returns

In this section we present a geometric approach to study the concept of risk adjusted returns. That is we quantify the excess (logarithmic) return an investor can achieve by augmenting the opportunity set and, at the same time, we provide geometric interpretations of Jensen's alpha and the beta parameter. While these quantities are considered fundamental for many portfolio managers the amount of information they carry is rather limited. In fact, as pointed out in Eqs. (13) and (14), the only information contained in Jensen's alpha is the sign; indicating whether to add a long or short infinitesimal position of an asset to an existing portfolio. Hence, the knowledge of alpha and beta is, by itself, not enough to determine how to form a portfolio that maximizes, say, the instantaneous Sharpe ratio or the rate of excess (logarithmic) return. The reason why alpha and beta fail to be self contained is due to the easily verifiable scaling properties

$$\alpha_0(\lambda_1\mathbf{w}_1, \lambda_2\mathbf{w}_2) = \lambda_1\alpha_0(\mathbf{w}_1,\mathbf{w}_2), \quad \beta_0(\lambda_1\mathbf{w}_1, \lambda_2\mathbf{w}_2) = \frac{\lambda_1}{\lambda_2}\beta_0(\mathbf{w}_1,\mathbf{w}_2). \tag{34}$$

In other words, given an arbitrary trading strategy, represented by $(\mathbf{w}_1, \mathbf{w}_2)$, we can apply leverage $(\lambda_1\mathbf{w}_1, \lambda_2\mathbf{w}_2)$ to achieve any targeted alpha and beta. Consequently, by only looking at the parameters alpha and beta we cannot distinguish, say, diversified portfolios from leveraged portfolios. Another reason why the risk adjusted excess return measure of Jensen is of little importance is because it ignores the event of loosing arbitrary large amount of money due to excessive leverage. We therefore suggest a slightly modified measure, called the risk adjusted Sharpe ratio, that carries more information.

In order to formulate our approach we first introduce some terminology. Given two subspaces U_1, U_2 with trivial intersection, $U_1 \cap U_2 = \{\mathbf{0}\}$, we let $U_1 \oplus U_2$ denote the direct sum and recall the similar concept for Hilbert spaces

$$\mathcal{H}_1 \oplus \mathcal{H}_2 = (U_1, \mathbf{V}_0[U_1^*]) \oplus (U_2, \mathbf{V}_0[U_2^*]) = (U_1 \oplus U_2, \mathbf{V}_0[U_1^*] \oplus \mathbf{V}_0[U_2^*]). \tag{35}$$

Hence, $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ if $U = U_1 \oplus U_2$ and $U_1 \perp U_2$. As mentioned in Remark 2 there is no real conceptual gain in explicitly expressing the space for which the inner product can be expanded in some basis. Consequently, from here and onward, we simply write $\mathbf{V}_0 \oplus \mathbf{V}_0$ unless there is ambiguity. We also write $\mathbf{w}_*$ when referring to $\mathbf{w}_*[U]$ and $U = U_1 \oplus U_2$. The following results show the importance of the Hilbert space direct sum decomposition.

Proposition 4.1. *Let $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. Then*

$$\begin{aligned}
\mathbf{w}_* &= \mathbf{w}_*[U_1] + \mathbf{w}_*[U_2], \\
\mathbf{s}_0^2(\mathbf{w}_*) &= \mathbf{s}_0^2(\mathbf{w}_*[U_1]) + \mathbf{s}_0^2(\mathbf{w}_*[U_2]), \\
\mathbf{b}_0(\mathbf{w}_*) &= \mathbf{b}_0(\mathbf{w}_*[U_1]) + \mathbf{b}_0(\mathbf{w}_*[U_2]).
\end{aligned}$$

Proof. Since $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ there is a unique decomposition $\mathbf{w}_* = \mathbf{w}_1 + \mathbf{w}_2$, such that $\mathbf{w}_i \in \mathcal{H}_i$. Because $U_1 \perp U_2$ we can further identify $\mathbf{w}_i$ with $\mathbf{P}_{0|U_i}(\mathbf{w}_*)$, from which the first result follows by Theorem 3.3. The second result also follows from Theorem 3.3, since $U_1 \perp U_2$, while the third result follows from the dual representation $\mathcal{H}^* = \mathcal{H}_1^* \oplus \mathcal{H}_2^*$. □

Corollary 4.2. *Let $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. Then*

$$\mathbf{b}_0(\mathbf{w}_*[U_i]) = \mathbf{s}_0^2(\mathbf{w}_*[U_i]), \quad \mu_0(\mathbf{w}_*[U_i]) = \frac{1}{2}\mathbf{s}_0^2(\mathbf{w}_*[U_i]), \quad i \in \{1, 2\}.$$

Proof. By the use of Lemma 3.2 and Theorem 3.3, it now follows that

$$\mathbf{b}_0(\mathbf{w}_*[U_i]) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{P}_{0|U_i}(\mathbf{w}_*)) = \mathbf{V}_0(\mathbf{P}_{0|U_i}(\mathbf{w}_*), \mathbf{P}_{0|U_i}(\mathbf{w}_*)) = \mathbf{s}_0^2(\mathbf{w}_*[U_i]),$$
$$\sigma_0^2(\mathbf{w}_*[U_i]) = \mathbf{V}_0(\mathbf{P}_{0|U_i}(\mathbf{w}_*), \mathbf{P}_{0|U_i}(\mathbf{w}_*)) = \mathbf{s}_0^2(\mathbf{w}_*[U_i]),$$

from which the proof follows. □

Hence, for a growth optimal Kelly trader the (logarithmic) excess return related to an augmentation of the opportunity set is directly linked to the Sharpe ratio. Note further that the key issue, as explained in Proposition 4.1 and Corollary 4.2, is to find the Hilbert space direct sums given two arbitrary (and thus not necessarily orthogonal) vector spaces $U_1, U_2 \subseteq U$. Henceforth, we let $U_{2|1}^{\perp}$ denote the orthogonal subspace to U_1 in U, while $U_{1|2}^{\perp}$ denotes the orthogonal subspace to U_2 in U, such that

$$\mathbf{P}_{0|U_{2|1}^{\perp}} = \mathbf{1}_U - \mathbf{P}_{0|U_1}, \quad \mathbf{P}_{0|U_{1|2}^{\perp}} = \mathbf{1}_U - \mathbf{P}_{0|U_2}. \tag{36}$$

This shows that $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_{2|1}^{\perp} = \mathcal{H}_{1|2}^{\perp} \oplus \mathcal{H}_2$, where

$$\mathcal{H}_1 \oplus \mathcal{H}_{2|1}^{\perp} = (U_1 \oplus U_{2|1}^{\perp}, \mathbf{V}_0 \oplus \mathbf{V}_0), \quad \mathcal{H}_{1|2}^{\perp} \oplus \mathcal{H}_2 = (U_{1|2}^{\perp} \oplus U_2, \mathbf{V}_0 \oplus \mathbf{V}_0). \tag{37}$$

The interpretation is that a growth optimal Kelly trader in U_1 should add the orthogonal vector $\mathbf{w}_*[U_{2|1}^{\perp}]$ to be growth optimal in U, while a growth optimal Kelly trader in U_2 should add the orthogonal vector $\mathbf{w}_*[U_{1|2}^{\perp}]$. In order to establish a connection to the alpha and beta parameters we further show that the instantaneous excess return covectors $\mathbf{b}_0[U_{2|1}^{\perp *}]$ and $\mathbf{b}_0[U_{1|2}^{\perp *}]$ are related to alpha, while the orthogonal projection operators $\mathbf{P}_{0|U_{2|1}^{\perp}}$ and $\mathbf{P}_{0|U_{1|2}^{\perp}}$ are linked to beta. We also stress that while the primary market might consist of, say, N numéraire based assets, we generally assume that only some mutual funds are available for investment. Consequently, $\dim(U) = \dim(U_1) + \dim(U_2) \leq N$. For ease of readability we choose to present our results in two steps: first we consider the simple case where $\dim(U_1) = \dim(U_2) = 1$ and thereafter we consider the general case. As always, most of the results carry over to higher dimensions albeit with some modifications.

4.1 Kelly Solution in Two Dimensions

Consider a market with only two investable assets such that $\dim(U) = 2$. We stress that each asset can be thought of as a mutual fund, with positions in a much larger asset universe. Let further $\mathbf{v}_1, \mathbf{v}_2 \in U$ be two linearly independent vectors (each corresponding to a particular trading strategy) and set $U_i = \mathrm{span}(\mathbf{v}_i)$, for $i = 1, 2$. Then, as $U_1 \cap U_2 = \{\mathbf{0}\}$, we have $U = U_1 \oplus U_2$. However, since the vectors $\mathbf{v}_1, \mathbf{v}_2$ are typically not orthogonal we cannot yet form the Hilbert space direct sum. For this reason we also consider the alternative decompositions $U = U_1 \oplus U_{2|1}^{\perp}$ and $U = U_{1|2}^{\perp} \oplus U_2$. While Corollary 4.2 formally identifies the risk adjusted quantities of interest we must always choose a particular basis for the computations. It should come as no surprise that these calculations can be greatly simplified if we use orthogonal basis vectors but that eventually we want to represent the risk adjustments using the natural basis vectors $(\mathbf{v}_1, \mathbf{v}_2)$. Hence, our first goal is to construct basis vectors $(\mathbf{v}_1, \mathbf{v}_{2|1})$ and $(\mathbf{v}_{1|2}, \mathbf{v}_2)$, where $\mathbf{v}_{2|1}$ is some vector spanning $U_{2|1}^{\perp}$ and similarly for $\mathbf{v}_{1|2}$. Below we show how to use the projection operators to construct three such sets of basis vectors and in doing so we derive a geometrical interpretation of alpha and beta.

Given the non-orthogonal natural basis $(\mathbf{v}_1, \mathbf{v}_2)$ we recall the canonical dual basis $(\mathbf{v}^1, \mathbf{v}^2)$ as introduced in Eq. (15). By the use of Lemma 3.2 we have

$$\mathbf{P}_{0|U_1} = \frac{\mathbf{V}_0(\mathbf{v}_1)}{\mathbf{V}_0(\mathbf{v}_1, \mathbf{v}_1)}\mathbf{v}_1 = \beta_0(\cdot, \mathbf{v}_1)\mathbf{v}_1, \quad \mathbf{P}_{0|U_2} = \frac{\mathbf{V}_0(\mathbf{v}_2)}{\mathbf{V}_0(\mathbf{v}_2, \mathbf{v}_2)}\mathbf{v}_2 = \beta_0(\cdot, \mathbf{v}_2)\mathbf{v}_2, \tag{38}$$

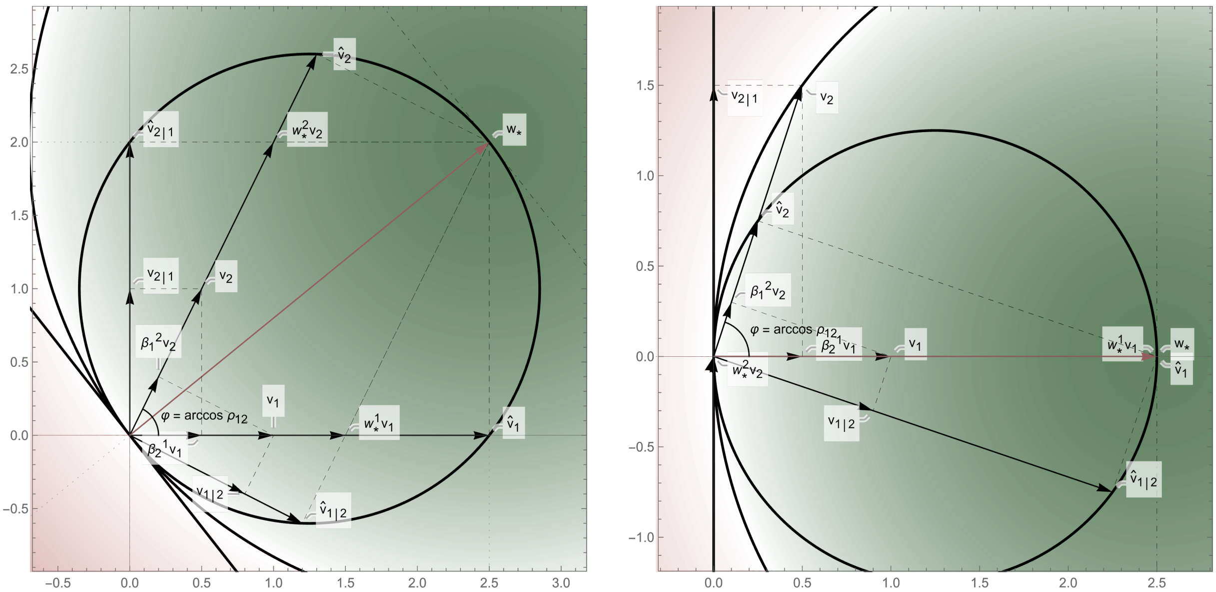
which identifies beta as being linked to the component of a projection tensor. Since the latter are (1,1)-tensors we can further expand them using the canonical dual basis and from Eq. (36) we get

$$\mathbf{P}_{0|U_{2|1}^{\perp}} = \sum_{i=1}^{2} \left(\mathbf{v}_i - \mathbf{P}_{0|U_1}(\mathbf{v}_i)\right) \mathbf{v}^i = \left(\mathbf{v}_2 - \beta_0(\mathbf{v}_2, \mathbf{v}_1)\mathbf{v}_1\right) \mathbf{v}^2, \tag{39}$$

$$\mathbf{P}_{0|U_{1|2}^{\perp}} = \sum_{i=1}^{2} \left(\mathbf{v}_i - \mathbf{P}_{0|U_2}(\mathbf{v}_i)\right) \mathbf{v}^i = \left(\mathbf{v}_1 - \beta_0(\mathbf{v}_1, \mathbf{v}_2)\mathbf{v}_2\right) \mathbf{v}^1. \tag{40}$$

We can now easily identify orthogonal vectors by setting

$$\mathbf{v}_{2|1} = \mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{v}_2) = \mathbf{v}_2 - \beta_0(\mathbf{v}_2, \mathbf{v}_1)\mathbf{v}_1, \quad \mathbf{v}_{1|2} = \mathbf{P}_{0|U_{1|2}^{\perp}}(\mathbf{v}_1) = \mathbf{v}_1 - \beta_0(\mathbf{v}_1, \mathbf{v}_2)\mathbf{v}_2. \tag{41}$$



(a) typical case (b) degenerate case

Figure 2: This figure shows the orthogonal decompositions $U = U_1 \oplus U_{2|1}^{\perp} = U_{1|2}^{\perp} \oplus U_2$ for two separate cases. The growth optimal Kelly vector $\mathbf{w}_* = \hat{\mathbf{v}}_1 + \hat{\mathbf{v}}_{2|1} = \hat{\mathbf{v}}_{1|2} + \hat{\mathbf{v}}_2$, which implies a representation $\mathbf{w}_* = w_*^1 \mathbf{v}_1 + w_*^2 \mathbf{v}_2$ in the non-orthogonal decomposition $U = U_1 \oplus U_2$. We use the notations: $\rho_{1,2} = \rho_0(\mathbf{v}_1, \mathbf{v}_2)$, ${\beta_1}^2 = \beta_0(\mathbf{v}_1, \mathbf{v}_2)$, ${\beta_2}^1 = \beta_0(\mathbf{v}_2, \mathbf{v}_1)$ and also highlight the level sets of $\mathbf{k}_0(\mathbf{w}) = k$, for $k \in \{1, 2, \pm\infty\}$.

In Fig. 2 we display the geometry of the orthogonal decompositions $U_1 \oplus U_{2|1}^{\perp}$ and $U_{1|2}^{\perp} \oplus U_2$, indicating the role of beta as being the components of a projection operator. From the degenerate case, plot (b), we further notice that, say, $\hat{\mathbf{v}}_{2|1} = \mathbf{w}_*[U_{2|1}^{\perp}] = \mathbf{0}$ is not equivalent to $\hat{\mathbf{v}}_2 = \mathbf{w}_*[U_2] = \mathbf{0}$. Having constructed the two auxiliary coordinate systems $(\mathbf{v}_1, \mathbf{v}_{2|1})$ and $(\mathbf{v}_{1|2}, \mathbf{v}_2)$ we proceed by investigating their local properties.

Proposition 4.3. *The characteristics of the vectors* $\mathbf{v}_{2|1}$ *and* $\mathbf{v}_{1|2}$ *are given by*

$$\mathbf{b}_0(\mathbf{v}_{2|1}) = \alpha_0(\mathbf{v}_2, \mathbf{v}_1), \quad \mathbf{V}_0(\mathbf{v}_{2|1}) = \sigma_0^2(\mathbf{v}_2) - \beta_0^2(\mathbf{v}_2, \mathbf{v}_1)\sigma_0^2(\mathbf{v}_1),$$
$$\mathbf{b}_0(\mathbf{v}_{1|2}) = \alpha_0(\mathbf{v}_1, \mathbf{v}_2), \quad \mathbf{V}_0(\mathbf{v}_{1|2}) = \sigma_0^2(\mathbf{v}_1) - \beta_0^2(\mathbf{v}_1, \mathbf{v}_2)\sigma_0^2(\mathbf{v}_2).$$

Proof. We illustrate the proof for the vector $\mathbf{w}_{2|1}$.

$$
\begin{aligned}
\mathbf{b}_0(\mathbf{v}_{2|1}) &= \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{2|1}) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_2) - \beta_0(\mathbf{v}_2, \mathbf{v}_1)\mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_1), \\
&= \mathbf{b}_0(\mathbf{v}_2) - \beta_0(\mathbf{v}_2, \mathbf{v}_1)\mathbf{b}_0(\mathbf{v}_1) = \alpha_0(\mathbf{v}_2, \mathbf{v}_1), \\
\mathbf{V}_0(\mathbf{v}_{2|1}, \mathbf{v}_{2|1}) &= \mathbf{V}_0(\mathbf{v}_2, \mathbf{v}_2) + \beta_0^2(\mathbf{v}_2, \mathbf{v}_1)\mathbf{V}_0(\mathbf{v}_1, \mathbf{v}_1) - 2\beta_0(\mathbf{v}_2, \mathbf{v}_1)\mathbf{V}_0(\mathbf{v}_1, \mathbf{v}_2), \\
&= \mathbf{V}_0(\mathbf{v}_2, \mathbf{v}_2) - \beta_0^2(\mathbf{v}_2, \mathbf{v}_1)\mathbf{V}_0(\mathbf{v}_1, \mathbf{v}_1).
\end{aligned}
$$

The proof for $\mathbf{w}_{1|2}$ is done analogously and is thus omitted. $\square$

It is important to notice that the knowledge of alpha and beta alone is not sufficient to calculate the corresponding Sharpe ratios $\mathbf{s}_0(\mathbf{v}_{2|1})$ and $\mathbf{s}_0(\mathbf{v}_{1|2})$, since these quantities also depend on the volatility of each trading strategy. We paraphrase this observation as:

Larger alpha with fixed beta is not necessarily better.

Smaller beta with fixed alpha is not necessarily better.

Furthermore, while it is pleasant to be able to interpret the excess return of the vectors $\mathbf{v}_{2|1}$ and $\mathbf{v}_{1|2}$ in terms of alpha, we must remember that the only purpose of these vectors is to span the spaces $U_{2|1}^{\perp}$ and $U_{1|2}^{\perp}$. Hence, any linear scaling of these vectors would serve equally well since the ultimate goal is to find $\hat{\mathbf{v}}_{2|1}$ and $\hat{\mathbf{v}}_{1|2}$. For this reason we rather prefer to express any risk adjustment in terms of Sharpe ratios as described below. This approach further reduces the number of free variables.

Definition 4.4. We call $\mathbf{s}_0(\mathbf{v}_{2|1})$ the *risk adjusted Sharpe ratio* of $\mathbf{v}_2$ given U_1 and define the corresponding risk adjustment of $\mathbf{v}_1$ given U_2 analogously.

Theorem 4.5. *The growth optimal Kelly vector admits the representation*

$$
\mathbf{w}_* = \frac{\mathbf{s}_0(\mathbf{v}_{1|2})}{\sigma_0(\mathbf{v}_1)\sqrt{1-\rho_0^2(\mathbf{v}_1, \mathbf{v}_2)}}\mathbf{v}_1 + \frac{\mathbf{s}_0(\mathbf{v}_{2|1})}{\sigma_0(\mathbf{v}_2)\sqrt{1-\rho_0^2(\mathbf{v}_1, \mathbf{v}_2)}}\mathbf{v}_2,
$$

where the instantaneous risk adjusted Sharpe ratios equal

$$
\mathbf{s}_0(\mathbf{v}_{2|1}) = \frac{\mathbf{s}_0(\mathbf{v}_2) - \rho_0(\mathbf{v}_1, \mathbf{v}_2)\mathbf{s}_0(\mathbf{v}_1)}{\sqrt{1-\rho_0^2(\mathbf{v}_1, \mathbf{v}_2)}}, \quad \mathbf{s}_0(\mathbf{v}_{1|2}) = \frac{\mathbf{s}_0(\mathbf{v}_1) - \rho_0(\mathbf{v}_1, \mathbf{v}_2)\mathbf{s}_0(\mathbf{v}_2)}{\sqrt{1-\rho_0^2(\mathbf{v}_1, \mathbf{v}_2)}}.
$$

Furthermore, the squared Sharpe ratio of the growth optimal Kelly strategy satisfy

$$\mathbf{s}_0^2(\mathbf{w}_*) = \mathbf{s}_0^2(\mathbf{v}_1) + \mathbf{s}_0^2(\mathbf{v}_{2|1}) = \mathbf{s}_0^2(\mathbf{v}_{1|2}) + \mathbf{s}_0^2(\mathbf{v}_2),$$
$$\mathbf{s}_0^2(\mathbf{w}_*) = \frac{\mathbf{s}_0^2(\mathbf{v}_1) + \mathbf{s}_0^2(\mathbf{v}_2) - 2\rho_0(\mathbf{v}_1, \mathbf{v}_2)\mathbf{s}_0(\mathbf{v}_1)\mathbf{s}_0(\mathbf{v}_2)}{1 - \rho_0^2(\mathbf{v}_1, \mathbf{v}_2)}.$$

Proof. By applying Proposition 4.1 and the notation in Eq. (26) we have

$$\mathbf{w}_* = \hat{\mathbf{v}}_1 + \hat{\mathbf{v}}_{2|1} = \frac{1}{\mathbf{k}_0(\mathbf{v}_1)}\mathbf{v}_1 + \frac{1}{\mathbf{k}_0(\mathbf{v}_{2|1})}\mathbf{v}_{2|1},$$
$$\mathbf{w}_* = \hat{\mathbf{v}}_{1|2} + \hat{\mathbf{v}}_2 = \frac{1}{\mathbf{k}_0(\mathbf{v}_{1|2})}\mathbf{v}_{1|2} + \frac{1}{\mathbf{k}_0(\mathbf{v}_2)}\mathbf{v}_2.$$

We now use Eq. (41) to transform these results to the coordinates $(\mathbf{v}_1, \mathbf{v}_2)$. Straightforward calculations yield

$$\mathbf{w}_* = \frac{1}{\mathbf{k}_0(\mathbf{v}_{1|2})}\mathbf{v}_1 + \frac{1}{\mathbf{k}_0(\mathbf{v}_{2|1})}\mathbf{v}_2.$$

The proof now follows from Propositions 4.1 and 4.3. □

The proof above uses the notion of relative leverage risk. Alternatively, one could express this variable in terms of the beta parameter (with respect to the growth optimal Kelly vector) as $1/\mathbf{k}_0(\mathbf{w}) = \beta_0(\mathbf{w}_*, \mathbf{w})$. However, since such a replacement brings little additional financial intuition, we leave this alternative as a curious observation. Instead, we provide two examples highlighting the behaviour in degenerate cases.

Example 4.6. Suppose that $\mathbf{s}_0(\mathbf{v}_2) = 0$, such that $\mathbf{s}_0(\mathbf{w}_*[U_2]) = 0$. Then, as shown in Theorem 3.3, $\|\mathbf{w}_*[U_2]\|_{\mathcal{H}} = 0$, or equally $\mathbf{w}_*[U_2] = \mathbf{0}$. But this does not imply that one should not invest in $\mathbf{v}_2$ when the opportunity set equals $U_1 \oplus U_2$. Rather, Theorem 4.5 gives

$$\mathbf{w}_* = \frac{\mathbf{s}_0(\mathbf{v}_1)}{\sigma_0(\mathbf{v}_1)(1 - \rho_0^2(\mathbf{v}_1, \mathbf{v}_2))}\mathbf{v}_1 - \frac{\rho_0(\mathbf{v}_1, \mathbf{v}_2)\mathbf{s}_0(\mathbf{v}_1)}{\sigma_0(\mathbf{v}_2)(1 - \rho_0^2(\mathbf{v}_1, \mathbf{v}_2))}\mathbf{v}_2.$$

Example 4.7. Suppose that $\mathbf{v}_2 = k\mathbf{w}_*$, such that $\mathbf{s}_0^2(\mathbf{v}_2) = \mathbf{s}_0^2(\mathbf{w}_*)$. Then, as shown in Theorem 4.5, $\mathbf{s}_0(\mathbf{v}_{1|2}) = 0$, or equally $\mathbf{s}_0(\mathbf{v}_1) = \rho_0(\mathbf{v}_1, \mathbf{v}_2)\mathbf{s}_0(\mathbf{v}_2)$. But this implies that

$$\mathbf{w}_* = \frac{\mathbf{s}_0(\mathbf{v}_{2|1})}{\sigma_0(\mathbf{v}_2)\sqrt{1 - \rho_0^2(\mathbf{v}_1, \mathbf{v}_2)}}\mathbf{v}_2 = \frac{\mathbf{s}_0(\mathbf{v}_2)}{\sigma_0(\mathbf{v}_2)}\mathbf{v}_2 = \mathbf{w}_*[U_2].$$

Hence, an investor who enlarges his opportunity set from U_1 to $U_1 \oplus U_2$ may well trade in the new asset even if it has zero Sharpe ratio. Moreover, such an investor may also fully discard his existing trading strategy in favour of only trading the asset in U_2 (even though the initial portfolio has non-zero Sharpe ratio).

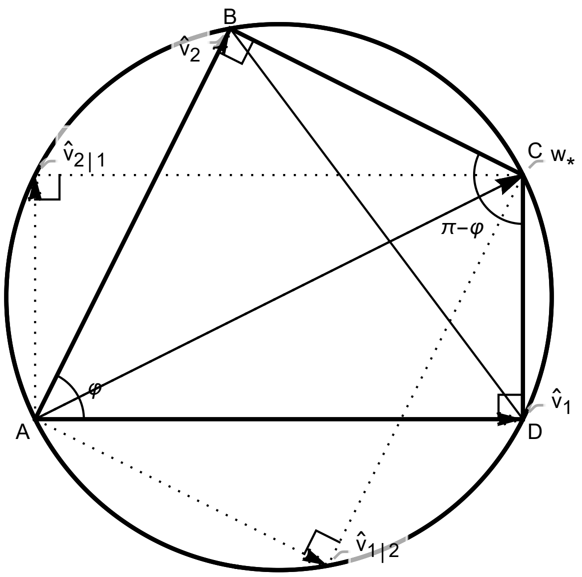


Figure 3: This figure shows that the orthogonal decompositions $U = U_1 \oplus U_{2|1}^{\perp} = U_{1|2}^{\perp} \oplus U_2$ form a cyclic quadrilateral. The circle, in which the quadrilateral is inscribed, corresponds to the level set of $\mathbf{k}_0(\mathbf{w}) = 1$, that is centered at $\mathbf{w}_*/2$ with a radius of $AC/2$. The quadrilateral is cyclic because opposite angles sum to π. Furthermore, the diagonals relate to the sides by Ptolemy's celebrated formula $BD \cdot AC = CD \cdot AB + AD \cdot BC$.

Finally, we take advantage of the 2-dimensional framework and present a pure geometric approach to identify the maximal Sharpe ratio and implicitly, thereby, the risk adjusted Sharpe ratios.

Example 4.8. From Fig. 3 and Ptolemy's formula, we know that

$$BD \cdot \|\mathbf{w}_*\|_{\mathcal{H}} = \|\hat{\mathbf{v}}_{2|1}\|_{\mathcal{H}} \|\hat{\mathbf{v}}_2\|_{\mathcal{H}} + \|\hat{\mathbf{v}}_1\|_{\mathcal{H}} \|\hat{\mathbf{v}}_{1|2}\|_{\mathcal{H}},$$

where BD represents the distance between the vectors $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$. We divide both sides with

$\|\mathbf{w}_*\|^2_{\mathcal{H}}$ and identify the ratios on the right hand side with angles according to

$$\frac{BD}{\|\mathbf{w}_*\|_{\mathcal{H}}} = \sin\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}_1}\cos\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}_2} + \cos\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}_1}\sin\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}_2} = \sin(\varphi_{\mathbf{w}_*,\hat{\mathbf{v}}_1} + \varphi_{\mathbf{w}_*,\hat{\mathbf{v}}_2}) = \sin\varphi_{\hat{\mathbf{v}}_1,\hat{\mathbf{v}}_2}.$$

It now follows from the law of cosines that

$$\|\mathbf{w}_*\|^2_{\mathcal{H}} = \frac{BD^2}{\sin^2\varphi_{\hat{\mathbf{v}}_1,\hat{\mathbf{v}}_2}} = \frac{\|\hat{\mathbf{v}}_1\|^2_{\mathcal{H}} + \|\hat{\mathbf{v}}_2\|^2_{\mathcal{H}} - 2\cos\varphi_{\hat{\mathbf{v}}_1,\hat{\mathbf{v}}_2}\|\hat{\mathbf{v}}_1\|_{\mathcal{H}}\|\hat{\mathbf{v}}_2\|_{\mathcal{H}}}{1-\cos^2\varphi_{\hat{\mathbf{v}}_1,\hat{\mathbf{v}}_2}},$$

which is the form presented in Theorem 4.5.

4.2 Kelly Solution in Arbitrary Dimensions

In this section we provide the solution of adding one opportunity set to another. The main difficulty lies in the fact that both opportunity sets consists of correlated assets; both among themselves but also among each other. In order to understand how the new assets affect the portfolio allocation we orthogonalize the covariance matrix, seen as a block matrix of the two sets of assets, in a way much similar to what was done in the previous section where the two sets only held one asset each. Although computing inverses of block matrices is well understood; using matrix formalism to solve our problem is, if not impossible, at least very difficult. We therefore adopt the formalism of geometric algebra and introduce, as before, three sets of intermingled coordinate systems.

In order to formulate the approach mathematically we consider two subspaces U_1 and U_2 of dimension N^1 and N^2, respectively. Each subspace is spanned by some linearly independent trading strategies and we use the notation $U_n = \operatorname{span}(\mathbf{v}_{1_n}, \ldots, \mathbf{v}_{N^n_n})$ to describe them. We further assume, without loss of generality, that $U_1 \cap U_2 = \{\mathbf{0}\}$ and form the direct sum $U = U_1 \oplus U_2$, such that $\dim(U) = N^1 + N^2$. For convenience we also introduce the notation

$$i_1 = i, \quad i_2 = N^1 + i, \tag{42}$$

such that we can identify the trading strategies in U when needed. Having defined our usage of multi-indices, we proceed by expanding the projection tensors, see Eq. (36), according to

$$\mathbf{P}_{0|U_{2|1}^{\perp}} = \sum_{n=1}^{2} \sum_{i_n=1_n}^{N_n^n} \left(\mathbf{v}_{i_n} - \mathbf{P}_{0|U_1}(\mathbf{v}_{i_n})\right) \mathbf{v}^{i_n} = \sum_{i_2=1_2}^{N_2^2} \left(\mathbf{v}_{i_2} - \mathbf{P}_{0|U_1}(\mathbf{v}_{i_2})\right) \mathbf{v}^{i_2}, \tag{43}$$

$$\mathbf{P}_{0|U_{1|2}^{\perp}} = \sum_{n=1}^{2} \sum_{i_n=1_n}^{N_n^n} \left(\mathbf{v}_{i_n} - \mathbf{P}_{0|U_2}(\mathbf{v}_{i_n})\right) \mathbf{v}^{i_n} = \sum_{i_1=1_1}^{N_1^1} \left(\mathbf{v}_{i_1} - \mathbf{P}_{0|U_2}(\mathbf{v}_{i_1})\right) \mathbf{v}^{i_1}, \tag{44}$$

where the canonical dual basis vectors satisfy $\mathbf{v}^{j_k}(\mathbf{v}_{i_l}) = \delta_{i_l}^{j_k}$. In order to further highlight the similarities with the 2-dimensional case, we use Einstein summation and write

$$\mathbf{P}_{0|U_1}(\mathbf{v}_{i_2}) = \beta_{i_2}{}^{k_1}\mathbf{v}_{k_1}, \quad \mathbf{P}_{0|U_2}(\mathbf{v}_{i_1}) = \beta_{i_1}{}^{k_2}\mathbf{v}_{k_2}, \tag{45}$$

in terms of some generalized beta parameters. Before we show how to compute the components of the projections we first introduce some notation.

Definition 4.9. For every subspace $\mathcal{H}_0 = (U_0, \mathbf{V}_0)$ of $\mathcal{H}$, let $\{\mathbf{V}_{0|U_0}^{i,j}\}$ denote the inverse of the Gram matrix on $\mathcal{H}_0 \subseteq \mathcal{H}$, such that for any chosen basis $\{\mathbf{v}_k\}_{k\leq K}$ of U_0, with $\dim(U_0) = K$, we have

$$\mathbf{V}_0(\mathbf{v}_i, \mathbf{v}_k)\mathbf{V}_{0|U_0}^{j,k} = \delta_i^j, \quad i, j \in \{1, \dots, K\},$$

Similarly, we let $\{\rho_{0|U_0}^{i,j}\}$ denote the inverse of the corresponding correlation matrix on $\mathcal{H}_0$, such that

$$\rho_{0|U_0}^{i,j} = \sigma_0(\mathbf{v}_i)\mathbf{V}_{0|U_0}^{i,j}\sigma_0(\mathbf{v}_j), \qquad \rho_0(\mathbf{v}_i, \mathbf{v}_k)\rho_{0|U_0}^{j,k} = \delta_i^j, \quad i, j \in \{1, \dots, K\}.$$

Lemma 4.10. *Let $\mathcal{H}_0 = (U_0, \mathbf{V}_0)$ be an arbitrary K-dimensional subspace of $\mathcal{H}$ and let $\{\mathbf{v}_k\}_{k\leq K}$ be a basis of U_0. Then, the projection*

$$\mathbf{P}_{0|U_0}(\mathbf{w}) = \mathbf{V}_0(\mathbf{w}, \mathbf{v}_j)\mathbf{V}_{0|U_0}^{j,k}\mathbf{v}_k,$$

of $\mathbf{w} \in \mathcal{H}$ onto $\mathcal{H}_0 \subseteq \mathcal{H}$ is orthogonal.

Proof. We first show that $\mathbf{P}_{0|U_0}$ is indeed a projection.

$$\begin{aligned}\mathbf{P}_{0|U_0}(\mathbf{P}_{0|U_0}(\mathbf{w})) &= \mathbf{V}_0(\mathbf{w},\mathbf{v}_j)\mathbf{V}_{0|U_0}^{j,k}\mathbf{P}_{0|U_0}(\mathbf{v}_k) = \mathbf{V}_0(\mathbf{w},\mathbf{v}_j)\mathbf{V}_{0|U_0}^{j,k}\mathbf{V}_0(\mathbf{v}_k,\mathbf{v}_a)\mathbf{V}_{0|U_0}^{a,b}\mathbf{v}_b,\\ &= \mathbf{V}_0(\mathbf{w},\mathbf{v}_j)\mathbf{V}_{0|U_0}^{j,k}\delta_k^b\mathbf{v}_b = \mathbf{V}_0(\mathbf{w},\mathbf{v}_j)\mathbf{V}_{0|U_0}^{j,k}\mathbf{v}_k = \mathbf{P}_{0|U_0}(\mathbf{w}).\end{aligned}$$

Next, we show that the projection is orthogonal

$$\begin{aligned}\mathbf{V}_0(\mathbf{P}_{0|U_0}(\mathbf{w}),\mathbf{P}_{0|U_0}(\mathbf{x})) &= \mathbf{V}_0(\mathbf{w},\mathbf{v}_j)\mathbf{V}_{0|U_0}^{j,k}\mathbf{V}_0(\mathbf{x},\mathbf{v}_a)\mathbf{V}_{0|U_0}^{a,b}\mathbf{V}_0(\mathbf{v}_k,\mathbf{v}_b),\\ &= \mathbf{V}_0(\mathbf{w},\mathbf{v}_j)\mathbf{V}_{0|U_0}^{j,k}\mathbf{V}_0(\mathbf{x},\mathbf{v}_a)\delta_k^a,\\ &= \mathbf{V}_0(\mathbf{w},\mathbf{v}_j)\mathbf{V}_{0|U_0}^{j,k}\mathbf{V}_0(\mathbf{x},\mathbf{v}_k) = \mathbf{V}_0(\mathbf{P}_{0|U_0}(\mathbf{w}),\mathbf{x}).\end{aligned}$$

Hence, $\mathbf{P}_{0|U_0}(\mathbf{w}) \perp \mathbf{x} - \mathbf{P}_{0|U_0}(\mathbf{x})$ which concludes the proof. □

We stress that the above result generalizes Lemma 3.2 by allowing the basis vectors to be non-orthogonal. Hence, rather than working with a non-observable abstract vector basis, we can directly consider the investable assets. Having identified the generalized beta parameters we now construct the vectors

$$\mathbf{v}_{i_2|1} = \mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{v}_{i_2}) = \mathbf{v}_{i_2} - \beta_{i_2}{}^{k_1}\mathbf{v}_{k_1}, \quad \beta_{i_2}{}^{k_1} = \mathbf{V}_0(\mathbf{v}_{i_2},\mathbf{v}_{j_1})\mathbf{V}_{0|U_1}^{j_1,k_1}, \tag{46}$$

$$\mathbf{v}_{i_1|2} = \mathbf{P}_{0|U_{1|2}^{\perp}}(\mathbf{v}_{i_1}) = \mathbf{v}_{i_1} - \beta_{i_1}{}^{k_2}\mathbf{v}_{k_2}, \quad \beta_{i_1}{}^{k_2} = \mathbf{V}_0(\mathbf{v}_{i_1},\mathbf{v}_{j_2})\mathbf{V}_{0|U_2}^{j_2,k_2}, \tag{47}$$

such that $U_1 \perp U_{2|1}^{\perp} = \mathrm{span}(\mathbf{v}_{1_2|1},\ldots,\mathbf{v}_{N_2^2|1})$ and $U_2 \perp U_{1|2}^{\perp} = \mathrm{span}(\mathbf{v}_{1_1|2},\ldots,\mathbf{v}_{N_1^1|2})$. The local properties of these orthogonal vectors are summarized below.

Proposition 4.11. *The characteristics of the vectors $\{\mathbf{v}_{i_2|1}\}$ and $\{\mathbf{v}_{i_1|2}\}$ are summarized by their instantaneous Sharpe ratios*

$$\begin{aligned}\mathbf{s}_0(\mathbf{v}_{i_2|1}) &= \frac{\mathbf{s}_0(\mathbf{v}_{i_2}) - \rho_0(\mathbf{v}_{i_2},\mathbf{v}_{j_1})\rho_{0|U_1}^{j_1,k_1}\mathbf{s}_0(\mathbf{v}_{k_1})}{\sqrt{1-\rho_0(\mathbf{v}_{i_2},\mathbf{v}_{j_1})\rho_{0|U_1}^{j_1,k_1}\rho_0(\mathbf{v}_{k_1},\mathbf{v}_{i_2})}},\\ \mathbf{s}_0(\mathbf{v}_{i_1|2}) &= \frac{\mathbf{s}_0(\mathbf{v}_{i_1}) - \rho_0(\mathbf{v}_{i_1},\mathbf{v}_{j_2})\rho_{0|U_2}^{j_2,k_2}\mathbf{s}_0(\mathbf{v}_{k_2})}{\sqrt{1-\rho_0(\mathbf{v}_{i_1},\mathbf{v}_{j_2})\rho_{0|U_2}^{j_2,k_2}\rho_0(\mathbf{v}_{k_2},\mathbf{v}_{i_1})}},\end{aligned}$$

and their instantaneous volatilities

$$\sigma_0(\mathbf{v}_{i_2|1}) = \sigma_0(\mathbf{v}_{i_2})\sqrt{1 - \rho_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_1})\rho_{0|U_1}^{j_1,k_1}\rho_0(\mathbf{v}_{k_1}, \mathbf{v}_{i_2})},$$
$$\sigma_0(\mathbf{v}_{i_1|2}) = \sigma_0(\mathbf{v}_{i_1})\sqrt{1 - \rho_0(\mathbf{v}_{i_1}, \mathbf{v}_{j_2})\rho_{0|U_2}^{j_2,k_2}\rho_0(\mathbf{v}_{k_2}, \mathbf{v}_{i_1})}.$$

Furthermore, the instantaneous correlation between the vectors, in each basis, equal

$$\rho_0(\mathbf{v}_{i_2|1}, \mathbf{v}_{j_2|1}) = \frac{\rho_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_2}) - \rho_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_1})\rho_{0|U_1}^{j_1,k_1}\rho_0(\mathbf{v}_{k_1}, \mathbf{v}_{j_2})}{\sqrt{1 - \rho_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_1})\rho_{0|U_1}^{j_1,k_1}\rho_0(\mathbf{v}_{k_1}, \mathbf{v}_{i_2})}\sqrt{1 - \rho_0(\mathbf{v}_{j_2}, \mathbf{v}_{j_1})\rho_{0|U_1}^{j_1,k_1}\rho_0(\mathbf{v}_{k_1}, \mathbf{v}_{j_2})}},$$
$$\rho_0(\mathbf{v}_{i_1|2}, \mathbf{v}_{j_1|2}) = \frac{\rho_0(\mathbf{v}_{i_1}, \mathbf{v}_{j_1}) - \rho_0(\mathbf{v}_{i_1}, \mathbf{v}_{j_2})\rho_{0|U_2}^{j_2,k_2}\rho_0(\mathbf{v}_{k_2}, \mathbf{v}_{j_1})}{\sqrt{1 - \rho_0(\mathbf{v}_{i_1}, \mathbf{v}_{j_2})\rho_{0|U_2}^{j_2,k_2}\rho_0(\mathbf{v}_{k_2}, \mathbf{v}_{i_1})}\sqrt{1 - \rho_0(\mathbf{v}_{j_1}, \mathbf{v}_{j_2})\rho_{0|U_2}^{j_2,k_2}\rho_0(\mathbf{v}_{k_2}, \mathbf{v}_{j_1})}}.$$

Proof. We only show how to compute the terms for $\mathbf{v}_{i_2|1}$ since $\mathbf{v}_{i_1|2}$ is treated similarly. First note that

$$\beta_{i_2}{}^{k_1}\mathbf{V}_0(\mathbf{v}_{k_1}, \mathbf{v}_{l_1}) = \mathbf{V}_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_1})\mathbf{V}_{0|U_1}^{j_1,k_1}\mathbf{V}_0(\mathbf{v}_{k_1}, \mathbf{v}_{l_1}) = \mathbf{V}_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_1})\delta_{l_1}^{j_1} = \mathbf{V}_0(\mathbf{v}_{i_2}, \mathbf{v}_{l_1}).$$

We therefore obtain

$$\begin{aligned}
\mathbf{V}_0(\mathbf{v}_{i_2|1}, \mathbf{v}_{j_2|1}) &= \mathbf{V}_0(\mathbf{v}_{i_2} - \beta_{i_2}{}^{k_1}\mathbf{v}_{k_1}, \mathbf{v}_{j_2} - \beta_{j_2}{}^{l_1}\mathbf{v}_{l_1}),\\
&= \mathbf{V}_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_2}) - \beta_{i_2}{}^{k_1}\mathbf{V}_0(\mathbf{v}_{k_1}, \mathbf{v}_{j_2}) - \beta_{j_2}{}^{l_1}\mathbf{V}_0(\mathbf{v}_{l_1}, \mathbf{v}_{i_2}) + \beta_{j_2}{}^{l_1}\beta_{i_2}{}^{k_1}\mathbf{V}_0(\mathbf{v}_{k_1}, \mathbf{v}_{l_1}),\\
&= \mathbf{V}_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_2}) - \beta_{i_2}{}^{k_1}\mathbf{V}_0(\mathbf{v}_{k_1}, \mathbf{v}_{j_2}),\\
&= \mathbf{V}_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_2}) - \mathbf{V}_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_1})\mathbf{V}_{0|U_1}^{j_1,k_1}\mathbf{V}_0(\mathbf{v}_{k_1}, \mathbf{v}_{j_2}).
\end{aligned}$$

We also calculate the generalized alpha representation

$$\mathbf{b}_0(\mathbf{v}_{i_2|1}) = \mathbf{b}_0(\mathbf{v}_{i_2}) - \beta_{i_2}{}^{k_1}\mathbf{b}_0(\mathbf{v}_{k_1}) = \mathbf{b}_0(\mathbf{v}_{i_2}) - \mathbf{V}_0(\mathbf{v}_{i_2}, \mathbf{v}_{j_1})\mathbf{V}_{0|U_1}^{j_1,k_1}\mathbf{b}_0(\mathbf{v}_{k_1}).$$

The proof concludes by replacing $\mathbf{V}_{0|U_1}$ by $\rho_{0|U_1}$, as described in Definition 4.9. □

It is of course a matter of taste which financial quantities to use when describing the local characteristics and here we deviate from Proposition 4.3 by focusing on risk adjusted Sharpe ratios, correlations and volatilities. The main reason for choosing these quantities is that the magnitude of both the Sharpe ratio and the correlation do not depend on leverage. The drawback is that neither quantity is a tensor, which means that sometimes it is easier to work with $(\mathbf{b}_0, \mathbf{V}_0)$. For those who ask for a geometric interpretation of the risk adjusted volatility we claim that

$$\cos^2 \varphi_{\mathbf{w},\mathbf{P}_{0|U_n}(\mathbf{w})} = \rho_0^2(\mathbf{w}, \mathbf{P}_{0|U_n}(\mathbf{w})) = \rho_0(\mathbf{w}, \mathbf{v}_{j_n})\rho_{0|U_n}^{j_n,k_n}\rho_0(\mathbf{v}_{k_n}, \mathbf{w}), \quad n \in \{1, 2\}. \tag{48}$$

The verification is left to the reader and instead we refer to Fig. 2 for a visualization in the 2-dimensional case. We now present the multi-dimensional extension of Theorem 4.5.

Theorem 4.12. *The growth optimal Kelly vector admits the representation*

$$\mathbf{w}_* = \mathbf{s}_0(\mathbf{v}_{i_1|2})\rho_{0|U_{1|2}^{\perp}}^{i_1,j_1}\sigma_0^{-1}(\mathbf{v}_{j_1|2})\mathbf{v}_{j_1} + \mathbf{s}_0(\mathbf{v}_{i_2|1})\rho_{0|U_{2|1}^{\perp}}^{i_2,j_2}\sigma_0^{-1}(\mathbf{v}_{j_2|1})\mathbf{v}_{j_2}.$$

Furthermore, the squared Sharpe ratio of the growth optimal Kelly strategy satisfy

$$\begin{aligned}
\mathbf{s}_0^2(\mathbf{w}_*) &= \mathbf{s}_0^2(\mathbf{w}_*[U_1]) + \mathbf{s}_0^2(\mathbf{w}_*[U_{2|1}^{\perp}]) = \mathbf{s}_0(\mathbf{v}_{i_1})\rho_{0|U_1}^{i_1,j_1}\mathbf{s}_0(\mathbf{v}_{j_1}) + \mathbf{s}_0(\mathbf{v}_{i_2|1})\rho_{0|U_{2|1}^{\perp}}^{i_2,j_2}\mathbf{s}_0(\mathbf{v}_{j_2|1}),\\
\mathbf{s}_0^2(\mathbf{w}_*) &= \mathbf{s}_0^2(\mathbf{w}_*[U_{1|2}^{\perp}]) + \mathbf{s}_0^2(\mathbf{w}_*[U_2]) = \mathbf{s}_0(\mathbf{v}_{i_1|2})\rho_{0|U_{1|2}^{\perp}}^{i_2,j_2}\mathbf{s}_0(\mathbf{v}_{j_1|2}) + \mathbf{s}_0(\mathbf{v}_{i_2})\rho_{0|U_2}^{i_2,j_2}\mathbf{s}_0(\mathbf{v}_{j_2}).
\end{aligned}$$

Proof. From Theorem 3.3 and Proposition 4.1 we have

$$\mathbf{w}_* = \mathbf{P}_{0|U_1}(\mathbf{w}_*) + \mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{w}_*) = \mathbf{P}_{0|U_{1|2}^{\perp}}(\mathbf{w}_*) + \mathbf{P}_{0|U_2}(\mathbf{w}_*).$$

Consequently, Lemma 4.10 gives us the equivalent expressions

$$\begin{aligned}
\mathbf{w}_* &= \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_1})\mathbf{V}_{0|U_1}^{i_1,j_1}\mathbf{v}_{j_1} + \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_2|1})\mathbf{V}_{0|U_{2|1}^{\perp}}^{i_2,j_2}\mathbf{v}_{j_2|1},\\
\mathbf{w}_* &= \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_1|2})\mathbf{V}_{0|U_{1|2}^{\perp}}^{i_1,j_1}\mathbf{v}_{j_1|2} + \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_2})\mathbf{V}_{0|U_2}^{i_2,j_2}\mathbf{v}_{j_2},
\end{aligned}$$

such that straightforward calculations yield

$$\begin{aligned}
\|\mathbf{w}_*\|_{\mathcal{H}}^2 &= \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_1})\mathbf{V}_{0|U_1}^{i_1,j_1}\mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{j_1}) + \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_2|1})\mathbf{V}_{0|U_{2|1}^{\perp}}^{i_2,j_2}\mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{j_2|1}),\\
\|\mathbf{w}_*\|_{\mathcal{H}}^2 &= \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_1|2})\mathbf{V}_{0|U_{1|2}^{\perp}}^{i_1,j_1}\mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{j_1|2}) + \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_2})\mathbf{V}_{0|U_2}^{i_2,j_2}\mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{j_2}).
\end{aligned}$$

Next, we use Eq. (46) to represent $\mathbf{w}_*$ in terms of the basis vectors $\{\mathbf{v}_{i_1}\}$ and $\{\mathbf{v}_{i_2}\}$. Similar to the proof of Theorem 4.5 we pick terms from each of the two representations to arrive at

$$\mathbf{w}_* = \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_1|2})\mathbf{V}_{0|U_{1|2}^{\perp}}^{i_1,j_1}\mathbf{v}_{j_1} + \mathbf{V}_0(\mathbf{w}_*, \mathbf{v}_{i_2|1})\mathbf{V}_{0|U_{2|1}^{\perp}}^{i_2,j_2}\mathbf{v}_{j_2}.$$

Finally, we use Definition 4.9 to express the results in terms of correlations rather than covariances. □

In order to verify that the above formula collapses to Theorem 4.5 when $N^1 = N^2 = 1$, we first notice that in this case $\mathbf{w}_* = \mathbf{k}_0^{-1}(\mathbf{v}_{1_1|2})\mathbf{v}_{1_1} + \mathbf{k}_0^{-1}(\mathbf{v}_{1_2|1})\mathbf{v}_{1_2}$. The correspondence then follows from converting the index references for each subspace U_1, U_2 to index references in U, as explained in Eq. (42). We continue with an example highlighting the benefits of diversification seen in higher dimensions. Loosely speaking we can think of the example as adding an asset to a trading strategy in, say S&P500, versus adding the asset to the opportunity set of the index.

Example 4.13. Let $N^1 > 1$ and $N^2 = 1$. In this example we study the difference in trading the assets $\{\mathbf{v}_{1_1}, \dots, \mathbf{v}_{N_1^1}, \mathbf{v}_{1_2}\}$ versus trading only in $\{\mathbf{w}_*[U_1], \mathbf{v}_{1_2}\}$. For sake of simplicity we introduce a new orthogonal basis $\{\check{\mathbf{e}}_{i_1}\}$ on $U_1 = \text{span}(\mathbf{v}_{1_1}, \dots, \mathbf{v}_{N_1^1})$, such that

$$\check{\mathbf{e}}_{1_1} = \frac{\mathbf{w}_*[U_1]}{\|\mathbf{w}_*[U_1]\|_{\mathcal{H}}}.$$

The inverse correlation matrix corresponding to the new Gram matrix on U_1 then takes the form $\check{\rho}_{0|U_1}^{j_1,k_1} = \delta^{j_1,k_1}$. Moreover, since $\mathbf{s}_0(\check{\mathbf{e}}_{i_1}) = \mathbf{s}_0(\mathbf{w}_*[U_1])$, if $i = 1$, and zero otherwise, Proposition 4.11 yields

$$\mathbf{s}_0^2(\mathbf{w}_*[U_{2|1}^{\perp}]) = \mathbf{s}_0^2(\mathbf{v}_{1_2|1}) = \frac{(\mathbf{s}_0(\mathbf{v}_{1_2}) - \rho_0(\mathbf{v}_{1_2}, \check{\mathbf{e}}_{1_1})\mathbf{s}_0(\check{\mathbf{e}}_{1_1}))^2}{1 - \sum_{i=1}^{N^1} \rho_0^2(\mathbf{v}_{1_2}, \check{\mathbf{e}}_{i_1})}.$$

We now compare this result with a growth optimal Kelly strategy on $\check{U} = \check{U}_1 \oplus U_2$, where $\check{U}_1 = \text{span}(\mathbf{w}_*[U_1]) = \text{span}(\check{\mathbf{e}}_{1_1})$. Consequently, we have $\mathbf{s}_0(\mathbf{w}_*[\check{U}_{2|1}^{\perp}]) = \mathbf{s}_0(\mathbf{w}_*[U_{2|1}^{\perp}])|_{N^1=1}$ and from Theorem 4.12 we calculate

$$\frac{\mathbf{s}_0^2(\mathbf{w}_*[U]) - \mathbf{s}_0^2(\mathbf{w}_*[\check{U}])}{(\mathbf{s}_0(\mathbf{v}_{1_2}) - \rho_0(\mathbf{v}_{1_2}, \mathbf{w}_*[U_1])\mathbf{s}_0(\mathbf{w}_*[U_1]))^2} = \frac{1}{1 - \sum_{i=1}^{N^1} \rho_0^2(\mathbf{v}_{1_2}, \check{\mathbf{e}}_{i_1})} - \frac{1}{1 - \rho_0^2(\mathbf{v}_{1_2}, \check{\mathbf{e}}_{1_1})} \geq 0,$$

with equality if and only if $\rho_0(\mathbf{v}_{1_2}, \check{\mathbf{e}}_{i_1}) = 0$, for $2_1 \leq i_1 \leq N_1^1$. In general, though, diversification has a positive effect on the maximal Sharpe ratio and thereby on the (logarithmic) excess return, for any Kelly trader.

Without going into details we mention that the previous example can easily be generalized to the situation where both $N^1, N^2 > 1$. In this case one finds that

$$\mathbf{s}_0^2(\mathbf{w}_*[U]) = \mathbf{s}_0^2(\mathbf{w}_*[\tilde{U}]), \quad \tilde{U} = \operatorname{span}(\mathbf{w}_*[U_1]) \oplus \operatorname{span}(\mathbf{w}_*[U_2]), \tag{49}$$

if and only if $\mathbf{w}_*[U_1] \perp \check{\mathbf{e}}_{2_2}, \ldots, \check{\mathbf{e}}_{N_2^2}$ and $\mathbf{w}_*[U_2] \perp \check{\mathbf{e}}_{2_1}, \ldots, \check{\mathbf{e}}_{N_1^1}$, where $\{\check{\mathbf{e}}_{i_k}\}_{i \geq 1}$ denotes an orthogonal basis in U_k, $k \in \{1, 2\}$, such that $\check{\mathbf{e}}_{1_k} = \mathbf{w}_*[U_k]/\|\mathbf{w}_*[U_k]\|_{\mathcal{H}}$.

We conclude this section by noting that Jensen's alpha, as a risk adjusted return, has a number of shortcomings. First, it does not specify the risk metric under which we can quantify excess return for a given level of risk. Second, it does not answer the question why its particular trading strategy is superior, or even preferable, to other trading strategies, and third, it does not readily generalize to higher dimensions since the diversification effect is not taken into account. In contrast, we argue that the Kelly approach, in combination with the risk adjusted Sharpe ratio, brings clarity to the picture and in the next section we provide further evidence supporting this claim.

5 Relative Value Trading

In this section we investigate the connection between relative value trading and option pricing as highlighted in Bermin and Holm (2021a). As shown in section 3.3, for a fixed level of logarithmic excess return, it is always favourable to use an efficient Kelly strategy; both in terms of relative leverage/drawdown risk and in terms of volatility. These properties follow from the fact that a Kelly strategy, by design, has maximal instantaneous Sharpe ratio and that it is never optimal to leverage more than the growth optimal Kelly strategy. We therefore choose to study the transformation of one efficient Kelly strategy to another as we enlarge the opportunity set. In other words we start with a trading strategy $\mathbf{w}_1 = k_1 \mathbf{w}_*[U_1]$, $k_1 \in [0, 1]$, and investigate the impact of extending the asset universe to $U = U_1 \oplus U_2$, when the new Kelly strategy $\mathbf{w} = k\mathbf{w}_*$ is used. From Eq. (32) we then have

$$\mu_0(\mathbf{w}_1) = \frac{1}{2}k_1(2-k_1)\mathbf{s}_0^2(\mathbf{w}_*[U_1]), \qquad \mu_0(\mathbf{w}) = \frac{1}{2}k(2-k)\mathbf{s}_0^2(\mathbf{w}_*),$$
$$\sigma_0^2(\mathbf{w}_1) = k_1^2\mathbf{s}_0^2(\mathbf{w}_*[U_1]), \qquad \sigma_0^2(\mathbf{w}) = k^2\mathbf{s}_0^2(\mathbf{w}_*),$$
$$\mathbf{k}_0(\mathbf{w}_1) = k_1, \qquad \mathbf{k}_0(\mathbf{w}) = k.$$

Hence, for a fixed relative leverage risk, $k = k_1$, the logarithmic excess return increases as

$$\mu_0(\mathbf{w}) - \mu_0(\mathbf{w}_1) = \frac{1}{2}k_1(2-k_1)\left(\mathbf{s}_0^2(\mathbf{w}_*) - \mathbf{s}_0^2(\mathbf{w}_*[U_1])\right) \geq 0. \tag{50}$$

If, instead, we keep the volatility fixed by setting $k = k_1\mathbf{s}_0(\mathbf{w}_*[U_1])/\mathbf{s}_0(\mathbf{w}_*)$, then

$$\mu_0(\mathbf{w}) - \mu_0(\mathbf{w}_1) = k_1\mathbf{s}_0(\mathbf{w}_*[U_1])\left(\mathbf{s}_0(\mathbf{w}_*) - \mathbf{s}_0(\mathbf{w}_*[U_1])\right) \geq 0. \tag{51}$$

Conversely, for a fixed logarithmic excess return we find that

$$k = 1 \pm \sqrt{1 - k_1(2-k_1)\frac{\mathbf{s}_0^2(\mathbf{w}_*[U_1])}{\mathbf{s}_0^2(\mathbf{w}_*)}}. \tag{52}$$

By choosing the efficient strategy with lowest variance (that is the one for which $k \in [0,1]$) we obtain after some algebraic manipulations

$$\frac{\mathbf{k}_0(\mathbf{w}) - \mathbf{k}_0(\mathbf{w}_1)}{1-k_1} = 1 - \sqrt{1 + \frac{k_1(2-k_1)}{(1-k_1)^2}\frac{(\mathbf{s}_0^2(\mathbf{w}_*) - \mathbf{s}_0^2(\mathbf{w}_*[U_1]))}{\mathbf{s}_0^2(\mathbf{w}_*)}} \leq 0, \tag{53}$$

$$\frac{\sigma_0(\mathbf{w}) - \sigma_0(\mathbf{w}_1)}{\mathbf{s}_0(\mathbf{w}_*) - k_1\mathbf{s}_0(\mathbf{w}_*[U_1])} = 1 - \sqrt{1 + 2k_1\frac{\mathbf{s}_0(\mathbf{w}_*[U_1])(\mathbf{s}_0(\mathbf{w}_*) - \mathbf{s}_0(\mathbf{w}_*[U_1]))}{(\mathbf{s}_0(\mathbf{w}_*) - k_1\mathbf{s}_0(\mathbf{w}_*[U_1]))^2}} \leq 0. \tag{54}$$

The conclusion to be drawn is that when restricted to trading strategies with maximal instantaneous Sharpe ratio it is almost always beneficial to enlarge the opportunity set. By doing so we can either increase the logarithmic excess return for a given relative leverage risk (or volatility) level or reduce the relative leverage risk (or volatility) for a given logarithmic excess return level. The only time when no value can be added, relative the initial portfolio, is when $\mathbf{s}_0(\mathbf{w}_*) = \mathbf{s}_0(\mathbf{w}_*[U_1])$. Here, the direct sum representation degenerates, see Fig. 2, and below we provide a number of equivalent conditions for when it happens.

Proposition 5.1. *Let $\mathcal{H} = (U_1 \oplus U_{2|1}^{\perp}, \mathbf{V}_0 \oplus \mathbf{V}_0)$. The following conditions are equivalent*

$$\begin{aligned} \mathbf{w}_* &= \mathbf{w}_*[U_1], & \mathbf{w}_*[U_{2|1}^{\perp}] &= \mathbf{0}, \\ \mathbf{b}_0(\mathbf{w}_*) &= \mathbf{b}_0(\mathbf{w}_*[U_1]), & \mathbf{b}_0(\mathbf{w}_*[U_{2|1}^{\perp}]) &= 0, \\ \mathbf{s}_0(\mathbf{w}_*) &= \mathbf{s}_0(\mathbf{w}_*[U_1]), & \mathbf{s}_0(\mathbf{w}_*[U_{2|1}^{\perp}]) &= 0, \end{aligned}$$

and

$$\begin{aligned} \mathbf{b}_0(\mathbf{w}) &= \mathbf{b}_0(\mathbf{P}_{0|U_1}(\mathbf{w})), \quad \forall \mathbf{w} \in \mathcal{H}, \\ \mathbf{s}_0(\mathbf{w}) &= \rho_0(\mathbf{w}, \mathbf{w}_*[U_1])\mathbf{s}_0(\mathbf{w}_*[U_1]), \quad \forall \mathbf{w} \in \mathcal{H}. \end{aligned}$$

Proof. Since $\mathbf{w}_* = \mathbf{w}_*[U_1] + \mathbf{w}_*[U_{2|1}^{\perp}]$, $\mathbf{b}_0(\mathbf{w}_*) = \mathbf{b}_0(\mathbf{w}_*[U_1]) + \mathbf{b}_0(\mathbf{w}_*[U_{2|1}^{\perp}])$, and $\mathbf{s}_0^2(\mathbf{w}_*) = \mathbf{s}_0^2(\mathbf{w}_*[U_1]) + \mathbf{s}_0^2(\mathbf{w}_*[U_{2|1}^{\perp}])$ the first set of conditions are trivially equivalent. Next, we prove that

$$\mathbf{s}_0^2(\mathbf{w}_*[U_{2|1}^{\perp}]) = 0 \Leftrightarrow \mathbf{b}_0(\mathbf{w}) = \mathbf{b}_0(\mathbf{P}_{0|U_1}(\mathbf{w})), \quad \forall \mathbf{w} \in \mathcal{H}.$$

We first note, since the projection operator $\mathbf{P}_{0|U_{2|1}^{\perp}}$ is orthogonal, that

$$\|\mathbf{w}_*[U_{2|1}^{\perp}]\|_{\mathcal{H}}^2 = \mathbf{V}_0(\mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{w}_*), \mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{w}_*)) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{w}_*)) = \mathbf{b}_0(\mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{w}_*)).$$

Hence, $\mathbf{s}_0^2(\mathbf{w}_*[U_{2|1}^{\perp}]) = \|\mathbf{w}_*[U_{2|1}^{\perp}]\|_{\mathcal{H}}^2 = 0$ if and only if the covector $\mathbf{b}_0 \circ \mathbf{P}_{0|U_{2|1}^{\perp}} = \mathbf{0}$. But this is equivalent to $\mathbf{b}_0(\mathbf{w}) = \mathbf{b}_0(\mathbf{P}_{0|U_1}(\mathbf{w}))$, for all $\mathbf{w} \in \mathcal{H}$, since

$$\mathbf{b}_0 \circ \mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{w}) = \mathbf{b}_0(\mathbf{P}_{0|U_{2|1}^{\perp}}(\mathbf{w})) = \mathbf{b}_0(\mathbf{w} - \mathbf{P}_{0|U_1}(\mathbf{w})) = \mathbf{b}_0(\mathbf{w}) - \mathbf{b}_0(\mathbf{P}_{0|U_1}(\mathbf{w})),$$

which proves the statement. Finally, we notice that

$$\mathbf{b}_0(\mathbf{w}) = \mathbf{b}_0(\mathbf{P}_{0|U_1}(\mathbf{w})) = \mathbf{V}_0(\mathbf{w}_*, \mathbf{P}_{0|U_1}(\mathbf{w})) = \mathbf{V}_0(\mathbf{P}_{0|U_1}(\mathbf{w}_*), \mathbf{w}),$$

is equivalent to

$$\mathbf{s}_0(\mathbf{w}) = \rho_0(\mathbf{P}_{0|U_1}(\mathbf{w}_*), \mathbf{w})\|\mathbf{P}_{0|U_1}(\mathbf{w}_*)\|_{\mathcal{H}},$$

from which the proof concludes by Theorem 3.3. □

Below we explain how the concept $\mathbf{s}_0(\mathbf{w}_*) = \mathbf{s}_0(\mathbf{w}_*[U_1])$ can be applied to the pricing of derivatives. We call the pricing rule No Added Relative Value (NARV, for short), with the meaning that the price of an asset is set such that there is no added value, relative to an existing Kelly portfolio, in trading the asset.

5.1 Derivative Pricing

In this section we explain how to price a derivative on one or several assets in a space U_1. Throughout this section, we let $\mathbf{v}_\pi$ denote a trading strategy that only takes positions in the derivative. As explained in the previous section a Kelly trader with opportunity set U_1 can add value to his portfolio by extending the opportunity set if $U_1 \cap \operatorname{span}(\mathbf{v}_\pi) = \{\mathbf{0}\}$ and $\mathbf{s}_0(\mathbf{w}_*[U_1 \oplus \operatorname{span}(\mathbf{v}_\pi)]) \neq \mathbf{s}_0(\mathbf{w}_*[U_1])$. Below, we analyze the meaning of these two conditions and highlight the connection with derivative pricing by means of no-arbitrage.

First, we observe that if $\mathbf{v}_\pi \in U_1$ then $U_1 \cap \operatorname{span}(\mathbf{v}_\pi) \neq \{\mathbf{0}\}$, with the interpretation that U_1 is instantaneously a complete market for valuing the derivative. From Theorem 3.5 we then have

$$\mathbf{s}_0(\mathbf{v}_\pi) = \rho_0(\mathbf{v}_\pi, \mathbf{w}_*[U_1])\mathbf{s}_0(\mathbf{w}_*[U_1]). \tag{55}$$

Note that when the derivative is written on one asset only, such that $\mathbf{v}_\pi = \lambda \mathbf{v}_{1_1}$, a repeated use of Theorem 3.5 verifies the well known expression $\mathbf{s}_0(\mathbf{v}_\pi) = \pm \mathbf{s}_0(\mathbf{v}_{1_1})$, with the sign depending on whether we are, for instance, considering a call or a put option. If we further require Eq. (55) to hold for each fixed point in time until the expiry of the derivative, the corresponding price is uniquely defined once we specify the terminal payoff of the derivative. We identify the price as the no-arbitrage price of Merton (1973), allowing for a synthetic replication of the terminal payoff by dynamically trading in the underlying assets.

Next, let us assume that $U_1 \cap \operatorname{span}(\mathbf{v}_\pi) = \{\mathbf{0}\}$, such that $\mathbf{v}_\pi \notin U_1$. In this case we say that U_1 is instantaneously an incomplete market with respect to the derivative. From Proposition 5.1 it then follows that the No Added Relative Value (NARV) price is characterized by

$$\mathbf{s}_0(\mathbf{w}_*[U_1 \oplus \operatorname{span}(\mathbf{v}_\pi)]) = \mathbf{s}_0(\mathbf{w}_*[U_1]) \Leftrightarrow \mathbf{s}_0(\mathbf{v}_\pi) = \rho_0(\mathbf{v}_\pi, \mathbf{w}_*[U_1])\mathbf{s}_0(\mathbf{w}_*[U_1]). \tag{56}$$

Hence, the local characteristics of the NARV price are identical to those of the no-arbitrage price in a complete market.

In order to further explain the properties of NARV pricing, we let $\mathbf{v}_\pi \in U_1 \oplus U_2$, for some set U_2. The interpretation is that $U_1 \oplus U_2$ is instantaneously a complete market or equally that the instantaneously incomplete market U_1 has been completed by adding the opportunity set U_2. The unique price of the derivative then satisfies $\mathbf{s}_0(\mathbf{v}_\pi) = \rho_0(\mathbf{v}_\pi, \mathbf{w}_*[U_1 \oplus U_2])\mathbf{s}_0(\mathbf{w}_*[U_1 \oplus U_2])$, as shown in Theorem 3.5. Consequently, the market completion adds no value, relative U_1, if $\mathbf{w}_*[U_1 \oplus U_2] = \mathbf{w}_*[U_1]$. But, as shown in Proposition 5.1, this is equivalent to

$$\mathbf{s}_0(\mathbf{w}) = \rho_0(\mathbf{w}, \mathbf{w}_*[U_1])\mathbf{s}_0(\mathbf{w}_*[U_1]), \quad \forall \mathbf{w} \in U_1 \oplus U_2. \tag{57}$$

Hence, in this case, the functional form of the local characteristics is similar for the derivative $\mathbf{w} = \mathbf{v}_\pi$ and for the assets $\mathbf{w} \in U_2$. In order to explain the significance of this observation let us consider a market exhibiting stochastic volatility. We assume that U_1 consists of only one asset and that we want to value, say, a call option with strike K_1. Moreover, we further assume that the price of the derivative (represented by $\mathbf{v}_{\pi_1}$) is uniquely defined once we augment the opportunity set with another call option (represented by $\mathbf{v}_{\pi_2}$) with, say, strike K_2. Then it is reasonable to claim, since we a priori do not know the price of either derivative, that it should not matter in which order we complete the market and this is exactly what NARV pricing achieves.

Another way to characterize the NARV prices is by recalling Theorem 3.1, where it was proved that in a complete market the market price of risk vector is identical to the growth optimal Kelly vector. Consequently, if $\mathbf{w}_*[U_1 \oplus U_2] = \mathbf{w}_*[U_1]$, Proposition 5.1 alternatively states that the NARV prices can be computed using a market price of risk process satisfying

$$\mathbf{\Theta}[U_1 \oplus U_2] = \mathbf{w}_*[U_1 \oplus U_2] = \mathbf{w}_*[U_1] \Rightarrow \mathbf{\Theta}[U_{2|1}^\perp] = \mathbf{0}, \tag{58}$$

for every fixed point in time. In the finance literature, the probability measure associated with such a market price of risk process is called the minimal martingale measure and was first introduced in Föllmer and Schweizer (1991). While the connection between Kelly trading and derivative pricing has been derived in Bermin and Holm (2021a), we believe our geometrical approach provides additional insights; notably by realizing that the market price of risk vector and the growth optimal Kelly vector are identical in a complete market.

Finally, we stress that should the market price of a derivative not equal the minimal martingale measure price, a Kelly trader can always add value to his portfolio by enlarging the opportunity set with the derivative, as explained in the first part of this section.

6 Comment on Risk Relativity

In this section we briefly outline how our framework can be extended to cover the situation where risk is measured relative an asset different from the numéraire. As an illustrative example we may consider a fund manager who benchmarks his performance against, say, bitcoin but reports his earnings in dollars. This leads us to develop a Kelly-like theory for hyperplanes, which are not necessarily going through origo and, hence, are not vector spaces but merely affine spaces. We proceed as follows: given that $U = \mathrm{span}(\mathbf{v}_1, \ldots, \mathbf{v}_N)$, we consider an K-dimensional hyperplane A, with $K \leq N$, defined such that for any point $\mathbf{w} \in A$ we can find coefficients $\{\lambda_i\}_{1 \leq i \leq K}$ satisfying $\mathbf{w} = \mathbf{v}_0 + \lambda^i(\mathbf{v}_i - \mathbf{v}_0)$, for some arbitrary point $\mathbf{v}_0$. With $\mathbf{u}$ denoting the reference vector two situations can now occur. Either $\mathbf{u}$ belongs to A or the reference vector lies outside of the hyperplane. In this paper we only consider the first case, which allows us to choose $\mathbf{v}_0 = \mathbf{u}$. It follows that we can can translate the hyperplane to origo by subtracting the reference vector and form the vector space $A_{\mathbf{u}} = A - \mathbf{u}$. We then define the Hilbert space $\mathcal{H}_{\mathbf{u}} = (A_{\mathbf{u}}, \mathbf{V}_{\mathbf{u}})$, where the inner product in $A_{\mathbf{u}}$ relates to that in U according to

$$\mathbf{V}_{\mathbf{u}}(\mathbf{v}_{\mathbf{u}}, \mathbf{w}_{\mathbf{u}}) = \mathbf{V}_0(\mathbf{v} - \mathbf{u}, \mathbf{w} - \mathbf{u}), \quad \mathbf{v}_{\mathbf{u}}, \mathbf{w}_{\mathbf{u}} \in A_{\mathbf{u}}. \tag{59}$$

One notes that the zero vector is the origo in each vector space U and $A_{\mathbf{u}}$, respectively, but when expressed in terms of U the origo of $A_{\mathbf{u}}$ equals the point associated with the reference vector. Following the notation in Bermin and Holm (2021b) we then define, in accordance with Proposition 2.1, the financial quantities

$$\mathbf{b}_{\mathbf{u}}(\mathbf{w}) = \mathbf{V}_0(\mathbf{w}_* - \mathbf{u}, \mathbf{w} - \mathbf{u}), \quad \sigma^2_{\mathbf{u}}(\mathbf{w}) = \mathbf{V}_0(\mathbf{w} - \mathbf{u}, \mathbf{w} - \mathbf{u}), \tag{60}$$

$$\rho_{\mathbf{u}}(\mathbf{v}, \mathbf{w}) = \frac{\mathbf{V}_0(\mathbf{v} - \mathbf{u}, \mathbf{w} - \mathbf{u})}{\sqrt{\mathbf{V}_0(\mathbf{v} - \mathbf{u}, \mathbf{v} - \mathbf{u})\mathbf{V}_0(\mathbf{w} - \mathbf{u}, \mathbf{w} - \mathbf{u})}}, \tag{61}$$

where, as usual, $\mathbf{w}_* = \mathbf{w}_*[U]$. We also set $\mu_{\mathbf{u}}(\mathbf{w}) = \mathbf{b}_{\mathbf{u}}(\mathbf{w}) - \frac{1}{2}\sigma^2_{\mathbf{u}}(\mathbf{w})$, $\mathbf{s}_{\mathbf{u}}(\mathbf{w}) = \mathbf{b}_{\mathbf{u}}(\mathbf{w})/\sigma_{\mathbf{u}}(\mathbf{w})$, and $\mathbf{k}_{\mathbf{u}}(\mathbf{w}) = \sigma_{\mathbf{u}}(\mathbf{w})/\mathbf{s}_{\mathbf{u}}(\mathbf{w})$. Note that while these definitions are natural from a financial point of view they come with the drawback that the tensor properties of $\mathbf{b}_0$ and $\mathbf{V}_0$ are lost. One way to overcome this issue would be to define $\mathbf{b}_{\mathbf{u}}(\mathbf{w}_{\mathbf{u}}) = \mathbf{V}_0(\mathbf{w}_* - \mathbf{u}, \mathbf{w}_{\mathbf{u}})$, similar to

$\mathbf{V_u}$ in Eq. (59). Hence, whichever notation that is most convenient to use might vary from application to application. With that being said, we continue by defining the growth optimal Kelly vector on the hyperplane $A = \mathbf{u} + A_\mathbf{u}$ according to

$$\mathbf{w}_*[A] = \arg\max_{\mathbf{w}\in A} \mu_\mathbf{u}(\mathbf{w}) = \mathbf{u} + \arg\max_{\mathbf{w_u}\in A_\mathbf{u}} \mu_\mathbf{u}(\mathbf{u}+\mathbf{w_u}), \tag{62}$$

such that $\mathbf{w}_*[A] = \mathbf{w}_*$ if $A_\mathbf{u}$ and U share the same point space. Similar to affine subspaces, we then define subspaces of the hyperplane A as being generated by subspaces of the corresponding vector space $A_\mathbf{u}$. Moreover, for any subspace $A_{\mathbf{u}1} \subseteq A_\mathbf{u}$ we let $\mathbf{P}_{\mathbf{u}|A_{\mathbf{u}1}}$ denote the orthogonal projection of $A_\mathbf{u}$ onto $A_{\mathbf{u}1}$, see Lemma 4.10 for related details. This allows us to generalize Theorems 3.3 and 3.5 as below.

Corollary 6.1. *For $\mathcal{H}_{\mathbf{u}1} = (A_{\mathbf{u}1}, \mathbf{V_u}) \subseteq \mathcal{H}_\mathbf{u}$, let $A_1 = \mathbf{u} + A_{\mathbf{u}1}$ be the associated subspace of A. Then*

$$\mathbf{w}_*[A_1] - \mathbf{u} = \mathbf{P}_{\mathbf{u}|A_{\mathbf{u}1}}(\mathbf{w}_* - \mathbf{u}), \quad \|\mathbf{w}_*[A_1] - \mathbf{u}\|_{\mathcal{H}_\mathbf{u}} = \mathbf{s_u}(\mathbf{w}_*[A_1]).$$

Proof. Straightforward calculations, setting $\mathbf{w}_{*\mathbf{u}} = \mathbf{w}_* - \mathbf{u}$ and assuming $\mathbf{w_u} \in A_{\mathbf{u}1}$, yield

$$\begin{aligned}
\mu_\mathbf{u}(\mathbf{u}+\mathbf{w_u}) =& \mathbf{V_u}(\mathbf{w}_{*\mathbf{u}}, \mathbf{w_u}) - \frac{1}{2}\mathbf{V_u}(\mathbf{w_u}, \mathbf{w_u}), \\
=& \mathbf{V_u}(\mathbf{P}_{\mathbf{u}|A_{\mathbf{u}1}}(\mathbf{w}_{*\mathbf{u}}), \mathbf{w_u}) - \frac{1}{2}\mathbf{V_u}(\mathbf{w_u}, \mathbf{w_u}), \\
=& \frac{1}{2}\|\mathbf{P}_{\mathbf{u}|A_{\mathbf{u}1}}(\mathbf{w}_{*\mathbf{u}})\|^2_{\mathcal{H}_\mathbf{u}} - \frac{1}{2}\|\mathbf{w_u} - \mathbf{P}_{\mathbf{u}|A_{\mathbf{u}1}}(\mathbf{w}_{*\mathbf{u}})\|^2_{\mathcal{H}_\mathbf{u}}.
\end{aligned}$$

Hence, $\arg\max_{\mathbf{w_u}\in A_{\mathbf{u}1}} \mu_\mathbf{u}(\mathbf{u}+\mathbf{w_u}) = \mathbf{P}_{\mathbf{u}|A_{\mathbf{u}1}}(\mathbf{w}_{*\mathbf{u}})$, from which the first part of the proof follows. The second part is a direct consequence of $\mathbf{P}_{\mathbf{u}|A_{\mathbf{u}1}}$ being an orthogonal projection. □

Corollary 6.2. *For $\mathcal{H}_{\mathbf{u}1} = (A_{\mathbf{u}1}, \mathbf{V_u}) \subseteq \mathcal{H}_\mathbf{u}$, let $A_1 = \mathbf{u} + A_{\mathbf{u}1}$ be the associated subspace of A. Then, for $\mathbf{v} \in A_1$, we have*

$$\rho_\mathbf{u}(\mathbf{v}, \mathbf{w}_*[A_1]) = \frac{\mathbf{s_u}(\mathbf{v})}{\mathbf{s_u}(\mathbf{w}_*[A_1])}.$$

Proof. The proof follows similarly to that of Theorem 3.5 and is thus omitted. □

In fact, all the results derived throughout this paper are presented in such a way that they can be modified by simply changing the reference vector. For instance, simple calculations,

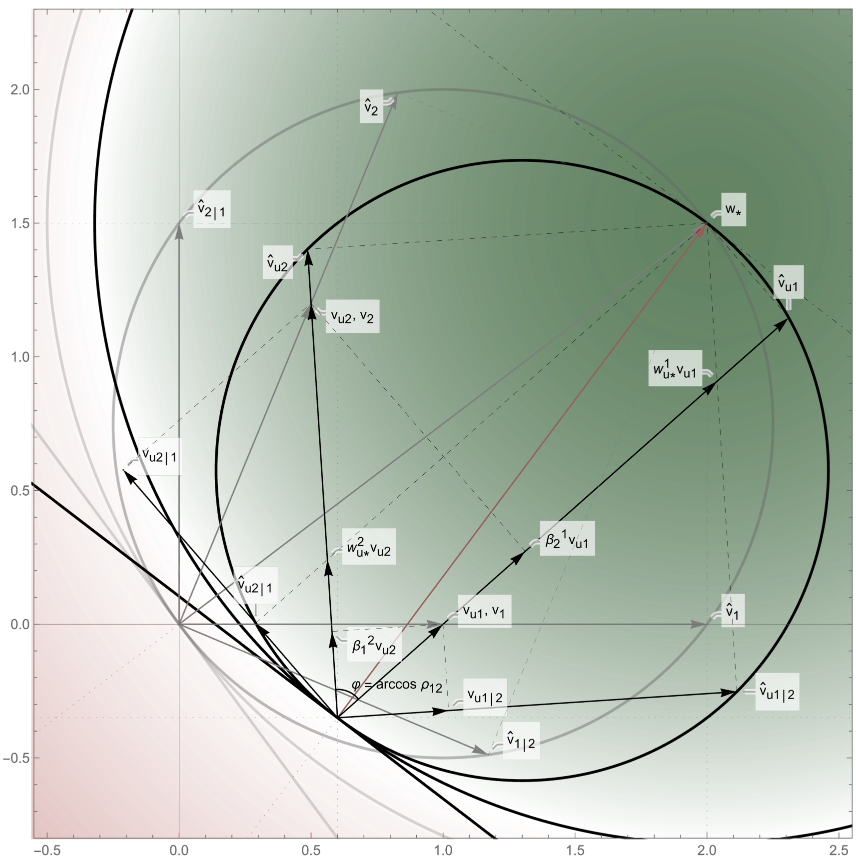


Figure 4: This figure shows the orthogonal decompositions of the translated vector space $A_{\mathbf{u}}$ (black) and those of the initial vector space U (grey). The growth optimal Kelly vector is invariant with respect to the translation vector $\mathbf{u}$, that is $\mathbf{w}_* = \mathbf{u} + \mathbf{w}_{\mathbf{u}*}$, which implies that Kelly strategies in $A_{\mathbf{u}}$ correspond to the trading strategies $\mathbf{w}_{\mathbf{u}} = k(\mathbf{w}_* - \mathbf{u})$. The growth optimal Kelly vector $\mathbf{w}_{\mathbf{u}*} = \widehat{\mathbf{v}_{\mathbf{u}1}} + \widehat{\mathbf{v}_{\mathbf{u}2|1}} = \widehat{\mathbf{v}_{\mathbf{u}1|2}} + \widehat{\mathbf{v}_{\mathbf{u}2}}$, further admits a representation $\mathbf{w}_{\mathbf{u}*} = w^1_{\mathbf{u}*}\mathbf{v}_{\mathbf{u}1} + w^2_{\mathbf{u}*}\mathbf{v}_{\mathbf{u}2}$ in the non-orthogonal decomposition $A_{\mathbf{u}} = A_{\mathbf{u}1} \oplus A_{\mathbf{u}2}$. We use the notations: $\rho_{1,2} = \rho_{\mathbf{u}}(\mathbf{v}_1, \mathbf{v}_2)$, $\beta_1{}^2 = \beta_{\mathbf{u}}(\mathbf{v}_1, \mathbf{v}_2)$, $\beta_2{}^1 = \beta_{\mathbf{u}}(\mathbf{v}_2, \mathbf{v}_1)$ and also highlight the level sets of $\mathbf{k}_{\mathbf{u}}(\mathbf{w}) = k$, for $k \in \{1, 2, \pm\infty\}$.

assuming $\dim(A_{\mathbf{u}}) = 2$, yield

$$\mathbf{s}^2_{\mathbf{u}}(\mathbf{w}_*) = \mathbf{s}^2_{\mathbf{u}}(\mathbf{v}_1) + \mathbf{s}^2_{\mathbf{u}}(\mathbf{v}_{2|1}) = \mathbf{s}^2_{\mathbf{u}}(\mathbf{v}_{1|2}) + \mathbf{s}^2_{\mathbf{u}}(\mathbf{v}_2), \tag{63}$$

where the risk adjusted Sharpe ratios equals

$$\mathbf{s}_{\mathbf{u}}(\mathbf{v}_{2|1}) = \frac{\mathbf{s}_{\mathbf{u}}(\mathbf{v}_2) - \rho_{\mathbf{u}}(\mathbf{v}_1, \mathbf{v}_2)\mathbf{s}_{\mathbf{u}}(\mathbf{v}_1)}{\sqrt{1 - \rho^2_{\mathbf{u}}(\mathbf{v}_1, \mathbf{v}_2)}}, \quad \mathbf{s}_{\mathbf{u}}(\mathbf{v}_{1|2}) = \frac{\mathbf{s}_{\mathbf{u}}(\mathbf{v}_1) - \rho_{\mathbf{u}}(\mathbf{v}_1, \mathbf{v}_2)\mathbf{s}_{\mathbf{u}}(\mathbf{v}_2)}{\sqrt{1 - \rho^2_{\mathbf{u}}(\mathbf{v}_1, \mathbf{v}_2)}}. \tag{64}$$

We visualise the role of what is considered to be the risk-free asset in Fig. 4. Although the lines $(A_{\mathbf{u}1}, A_{\mathbf{u}2})$, spanned by $\mathbf{u} + \lambda^1(\mathbf{v}_1 - \mathbf{u})$ and $\mathbf{u} + \lambda^2(\mathbf{v}_2 - \mathbf{u})$, respectively, are very different from the lines (U_1, U_2), spanned by $\lambda^1\mathbf{v}_1$ and $\lambda^2\mathbf{v}_2$, the direct sums $A_{\mathbf{u}1} \oplus A_{\mathbf{u}2}$ and $U_1 \oplus U_2$ have the same point space. Consequently, $\mathbf{w}_*[A] = \mathbf{w}_*$ and the trading strategies with maximal Sharpe ratio (i.e. the Kelly strategies) are now of the form $\mathbf{w} = \mathbf{u} + k(\mathbf{w}_* - \mathbf{u})$. For such trading strategies one easily verifies that

$$\mu_{\mathbf{u}}(\mathbf{w}) = \frac{1}{2}k(2-k)\mathbf{s}^2_{\mathbf{u}}(\mathbf{w}_*), \quad \sigma^2_{\mathbf{u}}(\mathbf{w}) = k^2\mathbf{s}^2_{\mathbf{u}}(\mathbf{w}_*), \quad \mathbf{k}_{\mathbf{u}}(\mathbf{w}) = k. \tag{65}$$

Hence, we recover the well known Kelly expressions. For higher dimensions Eq. (63) must, however, be modified as described in Theorem 4.12. We leave the details to the reader. Finally, we stress that the restricted growth optimal Kelly vectors $\mathbf{w}_*[A_1]$, $A_1 \subseteq A$, can change considerably with respect to the chosen reference vector $\mathbf{u}$, even though $\mathbf{w}_*[A]$ is invariant.

7 Conclusions

In this paper we present a geometric approach to portfolio theory, with the aim to explain and clarify the geometrical principles behind risk adjusted returns; in particular Jensen's alpha. We find that while the alpha/beta approach has severe limitations (especially in higher dimensions), only minor conceptual modifications are needed to complete the picture. However, these minor modifications (e.g. using risk adjusted Sharpe ratios rather than risk adjusted returns) can only be appreciated once a full geometric approach to portfolio theory is developed. In particular, we show how to create trading strategies on the efficient (local) frontier, in the sense of Markowitz (1952) and Tobin (1958), having maximal instantaneous Sharpe ratio. The approach taken is strongly linked to the Kelly criterion and the growth optimal Kelly vector.

Additionally, we derive a number of intermediate results that are of interest by themselves. For instance, we show that in a complete market the so called market price of risk vector is identical to the growth optimal Kelly vector, albeit expressed in coordinates of a different basis. We further show that the instantaneous correlation between an arbitrary trading strategy and its corresponding growth optimal Kelly strategy can be expressed as the ratio between their Sharpe ratios. By analyzing the level sets of various financial quantities we also find that points in the mean-variance space cannot, in general, be associated with a unique trading strategy. Only the points on the efficient frontier (that is those with maximal Sharpe ratio) can uniquely be identified. For such trading strategies, collinear to the growth optimal Kelly vector, we formalise the notion of relative value trading that is implicit in Platen (2006) and Bermin and Holm (2021a). We then apply geometric principles to investigate derivative pricing and introduce the concept of pricing by means on No Added Relative Value (NARV, for short). We say that this concept applies when the risk adjusted Sharpe ratio of the derivative equals zero. Using simple geometric arguments we show that NARV pricing is identical to no-arbitrage pricing with the so called minimal martingale measure of Föllmer and Schweizer (1991); a result first derived in Bermin and Holm (2021a), albeit with much different methods. We further show that should the market price of a derivative not equal the minimal martingale measure price, a Kelly trader can always add value to his portfolio by enlarging the opportunity set with the derivative.

References

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David D. Heath, 1999, Coherent Measures of Risk, *Mathematical Finance* 9, 203–228. https://doi.org/10.1111/1467-9965.00068

Bermin, Hans-Peter, and Magnus Holm, 2021a, Kelly trading and option pricing. *Journal of Futures Markets* 41, 987—1006. https://doi.org/10.1002/fut.22210

Bermin, Hans-Peter, and Magnus Holm, 2021b, Leverage and risk relativity: how to beat an index, KWC Working Paper 2019:2, Lund University. https://www.lusem.lu.se/media/kwc/working-papers/2019/kwcwp2019_2.pdf

Dodson, Christopher T.J., and Timothy Poston, 1991, *Tensor geometry* (Graduate Texts in Mathematics, 130, Springer-Verlag, Berlin, New York, Heidelberg). https://doi.org/10.1007/978-3-642-10514-2

Föllmer, Hans, and Alexander Schied, 2002, Convex measures of risk and trading constraints, *Finance and Stochastics* 6, 429–447. https://doi.org/10.1007/s007800200072

Föllmer, Hans, and Martin Schweizer, 1991, Hedging of contingent claims under incomplete information. in: Mark H.A. Davis, and Robert J. Elliott, ed.: *Applied Stochastic Analysis* (Stochastics Monographs 5, Gordon and Breach, New York)

Hakansson, Nils H., and William T. Ziemba, 1995, Capital Growth Theory, *Handbooks in Operations Research and Management Science* 9, 65–86. https://doi.org/10.1016/S0927-0507(05)80047-7.

Jensen, Michael C., 1964, The Performance of Mutual Funds in the Period 1945-1964. *Journal of Finance* 23, 389–416. https://doi.org/10.1111/j.1540-6261.1968.tb00815.x

Karatzas, Ioannis, and Steven S. Shreve, 1988, *Brownian Motion and Stochastic Calculus* (Springer-Verlag, New York Heidelberg Berlin).

Karatzas, Ioannis, and Steven S. Shreve, 1999, *Methods of Mathematical Finance* (Springer-Verlag, New York).

Kelly, John L., 1956, A new interpretation of information rate, *Bell System Technical Journal* 35, 917–926. https://doi.org/10.1002/j.1538-7305.1956.tb03809.x

Latané, Henry A., 1959, Criteria for choice among risky assets, *Journal of Political Economy* 67, 144–155. https://doi.org/10.1086/258157

Long, John B., 1990, The numeraire portfolio, *Journal of Financial Economics* 26. 29–69. https://doi.org/10.1016/0304-405X(90)90012-O

Luenberger, David G., 1997, *Optimization by Vector Space Methods* (John Wiley & Sons, New York).

MacLean, Leonard C., William T. Ziemba, and George Blazenko, 1992, Growth versus security in dynamic investment analysis, *Management Science* 38, 1562–1585. https://doi.org/10.1287/mnsc.38.11.1562

Markowitz, Harry, 1952, Portfolio Selection, *Journal of Finance* 7, 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Merton, Robert C., 1973, The Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science* 4, 141–183. https://doi.org/10.2307/3003143

Nielsen, Lars T., and Maria Vassalou, 2004, Sharpe Ratios and Alphas in Continuous Time, *Journal of Financial and Quantitative Analysis* 39, 103–114. https://doi.org/10.1017/S0022109000003902

Platen, Eckhard, 2006, A Benchmark Approach to Finance, *Mathematical Finance* 16, 131–151. https://doi.org/10.1111/j.1467-9965.2006.00265.x

Sharpe, William F., 1966, Mutual fund performance, *Journal of Business* 39, 119–138. http://doi.org/10.1086/294846

Sharpe, William F., 1994, The Sharpe Ratio, *Journal of Portfolio Management* 21, 49–58. http://doi.org/10.3905/jpm.1994.409501

Thorp, Edward O., 2011, The Kelly criterion in blackjack, sports betting and the stock market, in: Leonard C. MacLean, Edward O. Thorp, and William T. Ziemba, ed.: *The Kelly capital growth investment criterion* (World Scientific, Singapore). https://doi.org/10.1142/7598

Tobin, James, 1958, Liquidity Preference as Behavior Toward Risk, *Review of Economic Studies* 25, 65–86. https://doi.org/10.2307/2296205

Ziemba, William T., 2015, A response to Professor Paul A. Samuelson's objections to Kelly capital growth investing. *Journal of Portfolio Management* 42, 153–167. https://doi.org/10.3905/jpm.2015.42.1.153

The Geometry of Risk Adjustments

HANS-PETER BERMIN AND MAGNUS HOLM

In this paper we present a geometric approach to portfolio theory, with the aim to explain the geometrical principles behind risk adjusted returns; in particular Jensen's alpha. We find that while the alpha/beta approach has severe limitations (especially in higher dimensions), only minor conceptual modifications are needed to complete the picture. However, these modifications (e.g. using risk adjusted Sharpe ratios rather than returns) can only be appreciated once a full geometric approach to portfolio theory is developed. We further show that, in a complete market, the so called market price of risk vector is identical to the growth optimal Kelly vector, albeit expressed in coordinates of a different basis. For trading strategies collinear to the growth optimal Kelly vector, we formalise a notion of relative value trading based on the risk adjusted Sharpe ratio. As an application we show that a derivative having a risk adjusted Sharpe ratio of zero has a corresponding price given by the the minimal martingale measure.

Keywords: Jensen's alpha, Kelly criterion, market price of risk, option pricing, geometry

THE KNUT WICKSELL CENTRE FOR FINANCIAL STUDIES

The Knut Wicksell Centre for Financial Studies conducts cutting-edge research in financial economics and related academic disciplines. Established in 2011, the Centre is a collaboration between Lund University School of Economics and Management and the Research Institute of Industrial Economics (IFN) in Stockholm. The Centre supports research projects, arranges seminars, and organizes conferences. A key goal of the Centre is to foster interaction between academics, practitioners and students to better understand current topics related to financial markets.

LUND UNIVERSITY

SCHOOL OF ECONOMICS AND MANAGEMENT

LUND UNIVERSITY
SCHOOL OF ECONOMICS AND MANAGEMENT
Working paper 2021:2
The Knut Wicksell Centre for Financial Studies