

Active portfolio management

b-tu

October 1, 2024

The goal of this article is to present a focussed and subjective summary on the ideas covered in the book by [Grinold and Kahn \[1999\]](#) and the surrounding literature.

1 Notation

In this work, we adopt our own notation, which differs slightly from the set up presented by [Grinold and Kahn \[1999\]](#). For bookkeeping purposes, we now introduce some of the key notation that will be used throughout this work.

Financial investments. We let \mathbb{I} denote the space (or universe) of investments. An individual investment will be denoted by $\mathcal{I} \in \mathbb{I}$. Example of investments include both single assets and a weighted collection of assets, otherwise known as a portfolio. Note that there is a distinction between whether an asset is long or short. That is, a long position on a particular asset represents a different investment than the corresponding short position. A similar distinction also holds for assets held within a portfolio.

Finite number of assets. It is often the case that we are only interested in studying the space $\mathbb{I}_N \subset \mathbb{I}$, which is the space of investments spanned by a finite set of assets $\mathbb{A}_N = \{\mathcal{A}_1, \dots, \mathcal{A}_N\} \subset \mathbb{I}$, where N is the number of assets. Notably, we can identify each possible investment in this space with an N -dimensional weight vector,

$$\mathcal{I} \in \mathbb{I}_N \longleftrightarrow \mathbf{h} = (h^{(1)}, \dots, h^{(N)}) \in \mathbb{H}_N, \quad (1)$$

where the space of weights are constrained to have a signed L^1 -norm of less than one:

$$\mathbb{H}_N = \left\{ \mathbf{h} \in \mathbb{R}^N : \sum_{n=1}^N h^{(n)} \leq 1 \right\}. \quad (2)$$

Conceptually, each component $h^{(n)} \in \mathbb{R}$ controls the weight associated with the asset \mathcal{A}_n for $n \in [N] := \{1, \dots, N\}$. These weight components can either be positive (for a long position) or negative (for a short position). When the signed L^1 -norm is one, then we say that we are fully invested.

2 Background concepts

The goal of this section is to highlight some of the core conventions and ideas that are often present in the finance literature.

2.1 Returns

The return of an investment is often reported in the percentage range. That is, one does not report the raw numerical return (i.e. the change in profit), but instead one reports the rate of return (i.e. the proportional change in profit). Concretely, suppose we have an investment $\mathcal{I} \in \mathbb{I}$ whose value is given by $v_{t_0}(\mathcal{I}) \in \mathbb{R}$ at time t_0 and $v_T(\mathcal{I}) \in \mathbb{R}$ at some future time t , then the return of this investment at this latest time is given by

$$r_t(\mathcal{I}) = \frac{v_t(\mathcal{I}) - v_{t_0}(\mathcal{I})}{v_{t_0}(\mathcal{I})} \in \mathbb{R}. \quad (3)$$

Notice that this number tends to infinity when the initial value approaches zero (i.e. in the limit we are getting infinite return from a prior investment of nothing).

Remark 2.1 (Other value streams) *In practice, one must also account for other profits gained by holding onto the investment in equation (3). For example, one might also include gains from dividends. Out of convenience, we ignore subtle technicalities such as this.*

2.2 Everything is relative

Performance assessment is traditionally defined on a relative basis. That is, one typically computes the performance of an investment in comparison to another one. More precisely, we are often interested in the relative return of an investment,

$$R_t(\mathcal{I}, \mathcal{B}) := r_t(\mathcal{I}) - r_t(\mathcal{B}), \quad (4)$$

where $r_t(\mathcal{I})$ and $r_t(\mathcal{B})$ denotes the return from an investment $\mathcal{I} \in \mathbb{I}$ and benchmark $\mathcal{B} \in \mathbb{I}$, at time $t > 0$, respectively. Practitioners often set this benchmark rate to either the risk-free rate or the market rate. Typically, when the risk-free rate is used, then the relative return is referred to as the excess return. Otherwise, when this benchmark is set more generally, then the corresponding relative return is referred to as the active return.

Risk-free rate. The risk-free rate, $r_t(\mathcal{B}^{\text{free}})$, refers to the (hypothetical) minimum rate of return that one might desire from an investment. Typically, this rate is set to the return of a financial product whose risk is minimal¹. In practice, it is common to see this rate being set to the return of some Treasury bill (T-bill) or a long-term government bond yield.

Market rate. The market rate, $r_t(\mathcal{B}^{\text{market}})$, usually refers to the return achieved by a financial product that tries to match the market as a whole (or at least some subset of it). This market portfolio is usually constructed by taking all of the assets in the desired market and then weighting them according to their market capitalisation.

Remark 2.2 (Empty investment) *Note that we recover the original notion of return (3) if we set our benchmark to the empty investment.*

2.3 The meaning of risk

Following Grinold and Kahn [1999], the risk associated with an investment is defined as the standard deviation of its return. More specifically, the risk associated with an investment $\mathcal{I} \in \mathbb{I}$, at a time $t > 0$, is given by

$$\text{Risk}[\mathcal{I}] := \text{Std}[r_t(\mathcal{I})] = \sqrt{\text{Var}[r_t(\mathcal{I})]}. \quad (5)$$

¹Ideally, one would like to consider a zero risk investment—although realistically there is always a risk associated with any investment.

Similarly, the risk of the active return (4) is given by

$$\text{Risk}[(\mathcal{I}, \mathcal{B})] := \text{Std}[R_t(\mathcal{I}, \mathcal{B})] = \sqrt{\text{Var}[R_t(\mathcal{I}, \mathcal{B})]}. \quad (6)$$

This latter risk is sometimes referred to as the active risk or the tracking error.

Other measures of risk. As discussed by Grinold and Kahn [1999, Chapter 3], there are many possible alternatives to the standard deviation as a notion of risk. Although empirically, they suggest that there does not appear to be much to gain from exploring these alternatives—at least from a performance-to-cost standpoint.

Idiosyncratic and systematic risk. The risk associated entirely from an individual asset $\mathcal{I} \in \mathbb{I}$ is referred to as the idiosyncratic risk of \mathcal{I} . These risk can typically be mitigated (or even eliminated) by constructing a more diversified portfolio. The other sources of risk are collectively referred to as the systematic risk of \mathbb{I} . Conceptually, these are the risk involved with holding any investment. By construction, systematic risk cannot be diversified away.

2.4 The nature of time

The distinction between an ex-ante (before the event) approach and an ex-post (after the event) approach is often very subtle in the finance literature. For example, one might take an ex-ante approach when one is interested in forecasting and making predictions of the future. In contrast, one might take an ex-post approach in the performance assessment setting, where one is interested in assessing the quality² of a strategy or a portfolio manager.

Risk analysis. A core problem in asset management is concerned with the quantification of risk. It is important to note that one is interested in estimating risk in both the ex-ante and ex-post setting. In the ex-ante setting, we are interested estimating risk in order to incorporate it into our active decision-making. Whilst, in the ex-post setting, we are interested in assessing risk in order to identify its main active drivers and also to properly assess the realised performance.

Information set. We denote the historical information set, at time $t > 0$, by $\mathbf{X}_t \in \mathbb{X}$, where \mathbb{X} denotes the set of all possible historical information sets. Confusingly, some authors abuse the notation of the historical information set and allow it to be defined ambiguously in order to accommodate for both the ex-ante and ex-post setting. More specifically, in the ex-ante (or ex-post) setting, these authors let \mathbf{X}_t denote the information set containing all of the information available at the beginning (or end) of time t , respectively. This convention is convenient from a writing perspective as it covers both of these temporal settings simultaneously. On the other hand, it is definitely confusing from a reader perspective who could easily be misled. Despite this weakness, we will also follow this convention as it allows for easier comparisons between this document and other related works.

2.5 Garbage in, garbage out

A central tenet that is echoed throughout the book by Grinold and Kahn [1999] is the idea that the ability to identify and exploit “superior information” is what separates a good active manager from a bad one. In other words, to generate excess returns, one needs to be diligent

²Here the term quality is defined in the philosophical sense. In general, it is very hard to actually quantify what distinguishes a skillful investor and a lucky one.

in both their ability to collect useful information and their ability to use it effectively. Most of the work presented by Grinold and Kahn [1999] focusses on addressing the latter problem. We will cover a selection of these ideas in the following sections [TODO:link](#). The former problem, however, is a practical challenge that is much more an art than a science. Regardless of how one approaches this information retrieval problem, one inevitably will fall upon the well-know adage: “garbage in, garbage out”. That is, our inability to identify information of a ‘reasonable’ quality will, more often than not, lead to useless output. Therefore, it remains an important and unquestionable part of asset management, that one must aspire to collect high quality data in order to facilitate better decision making.

2.6 The efficient market hypothesis

The efficient-market hypothesis (EMH) states that the price as an asset or investment reflects all of the available information. That is, there should not be consistent opportunities for an investor to outperform the market.

Other variants. There are a number of different variants to the EMH. At a theoretical level, these variants mainly differ on what the concept of “available information” means. Namely, whether this is just any data available about past prices (weak), any publicly available information (semi-strong) or any possible information that exist in the world (strong).

Validity. The debate on whether the EMH (or some variant of it) holds in practice is contentious. [TODO:ref something](#).

2.7 Mean-variance trade-off

The goal of Modern Portfolio Theory (MPT) [TODO: cite Markowitz](#) is to find portfolios which achieves the greatest expected return for a given user-specified risk level. More formally, suppose we are working in the space \mathbb{I}_N , then for a given risk level $c \geq 0$, the goal of interest is to solve the following mean-variance optimisation problem:

$$\max_{\mathbf{h} \in \mathbb{I}_N : \mathbf{h}^T \boldsymbol{\Sigma}_t \mathbf{h} \leq c^2} \mathbf{h}^T \boldsymbol{\mu}_t, \quad (7)$$

for any time $t > 0$, where the $\boldsymbol{\mu}_t \in \mathbb{R}^N$ and $\boldsymbol{\Sigma}_t \in \mathbb{R}^{N \times N}$ denotes the mean and covariance matrix of the excess returns of the base assets,

$$\mathbf{R}_t := (R_t(\mathcal{A}_1, \mathcal{B}^{\text{free}}), \dots, R_t(\mathcal{A}_N, \mathcal{B}^{\text{free}}))^T \in \mathbb{R}^N, \quad (8)$$

$$\boldsymbol{\mu}_t := \mathbb{E}[\mathbf{R}_t] \in \mathbb{R}^N, \quad (9)$$

$$\boldsymbol{\Sigma}_t := \text{Var}[\mathbf{R}_t] \in \mathbb{R}^{N \times N}. \quad (10)$$

Notably, one of the key challenges involved with deploying an MPT solution comes from the fact that the mean vector and covariance matrix of the excess returns vector is unknown in practice and therefore has to be estimated (or predicted) from the available data. The mean vector has N elements that must be estimated, whilst the covariance matrix has $\mathcal{O}(N^2)$ number of elements to be estimated. In practice, one does not usually have enough data points at each time t in order to efficiently estimate the covariance matrix with high precision. As a result, practitioners typically try to reduce the dimensionality of this estimation problem by making some reasonable modelling assumptions on the excess returns of the assets. For instance a popular modelling framework are factor models (Section 3.1), which effectively reduces the dimensionality of the covariance matrix to $\mathcal{O}(K^2 + N)$, where K is the number of factors—see Section 3.1 for more details.

3 Modelling the returns

One of the main problems in active portfolio management is concerned with the modelling of excess returns (4). In particular, one often tries to model the excess returns of an investment $\mathcal{I} \in \mathbb{I}$, at time $t > 0$, as a function of the information set:

$$R_t(\mathcal{I}, \mathcal{B}^{\text{free}}) = f_t(\mathbf{X}_t) + \epsilon_t, \quad (11)$$

where $\mathbf{X}_t \in \mathbb{X}$ denotes the t -th information set, $f_t : \mathbb{X} \rightarrow \mathbb{R}$ denotes the (potentially random) modelling function and $\epsilon_t \in \mathbb{R}$ denotes the random residual. In the following paragraphs, we list some notable ideas that underpin this general model for excess returns (11).

Dependence on the investor. Notably, the form of the information set \mathbf{X}_t and function f_t is something that is specified by the user; the residual ϵ_t denotes everything else that is left over. On practical level, this means that different investors can obtain different results even if they adopt the same modelling function f_t .

Information uncertainty. In the most general case, the information set \mathbf{X}_t can be treated as a random variable. The randomness of this quantity stems from the intrinsic uncertainty that naturally arises from the collection, retrieval and processing of real-world data. Conventionally though, it is common to see this information set being treated as a deterministic realisation of some estimation or computation procedure.

Residual uncertainty. The residual $\epsilon_t = R_t(\mathcal{I}, \mathcal{B}^{\text{free}}) - f_t(\mathbf{X}_t)$ models the remaining discrepancy that is not accounted for by the “explained” part of the excess returns $f_t(\mathbf{X}_t)$. Note here that the residual ϵ_t is a random variable because the excess returns $R_t(\mathcal{I}, \mathcal{B}^{\text{free}})$ is a random variable.

Space of models. The modelling function $f_t : \mathbb{X} \rightarrow \mathbb{R}$ is assumed to live in some space \mathbb{M} . For example, much of the earlier work in the space assumed that \mathbb{M} was the space of linear functions. More recently, there has been a growing interest in using machine learning methods in order to learn more expressive non-linear models.

3.1 Factor models

Early work on modelling returns centred on the setting where the space of models \mathbb{M} was the space of linear models. The corresponding information sets \mathbb{X}_t , at each time $t > 0$, are then given by matrices whose columns represent any of the features that we think are relevant to modelling the excess returns. Formally speaking, these features are commonly referred to as the factor exposures (or factor loadings) of our investment, whilst the resulting linear models are referred to as the factor models. We now present a targetted overview of factor models. Following the example of existing work, we focus our of attention on the realistic setting where our the space of investments is \mathbb{I}_N .

3.1.1 Multi-factor model

Given an investment $\mathcal{I} \in \mathbb{I}_N$, represented by a weight vector $\mathbf{h} \in \mathbb{H}_N$, the multi-factor model assumes that the excess returns obeys the linear model

$$R_t(\mathcal{I}, \mathcal{B}^{\text{free}}) = \mathbf{h}^T (\mathbf{X}_t \mathbf{b}_t + \mathbf{u}_t) \quad (12)$$

$$= \sum_{k=1}^K \left(\sum_{n=1}^N h^{(n)} X_t^{(n,k)} \right) b_t^{(k)} + \sum_{n=1}^N h^{(n)} u_t^{(n)}, \quad (13)$$

at time $t > 0$, where:

- $\mathbf{X}_t \in \mathbb{R}^{N \times K}$ is the exposure matrix, whose elements $X_t^{(n,k)} \in \mathbb{R}$ are the factor loadings which represent the exposure (or sensitivity) of asset $n \in [N]$ to factor $k \in [K]$. This is commonly treated as a deterministic variable.
- $\mathbf{b}_t = (b_t^{(1)}, \dots, b_t^{(K)})^T \in \mathbb{R}^K$ is the factor returns vector, which is treated as a random variable.
- $\mathbf{u}_t = (u_t^{(1)}, \dots, u_t^{(N)})^T \in \mathbb{R}^N$ are the asset residuals, or idiosyncratic returns, which describe the excess returns of an asset that cannot be explained by the (systematic) factors.

The expectation and variance of (13) is then given by:

$$\mathbb{E}[R_t(\mathcal{I}, \mathcal{B}^{\text{free}})] = \mathbf{h}^T (\mathbb{E}[\mathbf{X}_t \mathbf{b}_t] + \mathbb{E}[\mathbf{u}_t]), \quad (14)$$

$$\text{Var}[R_t(\mathcal{I}, \mathcal{B}^{\text{free}})] = \mathbf{h}^T \text{Var}[\mathbf{X}_t \mathbf{b}_t + \mathbf{u}_t] \mathbf{h} \quad (15)$$

$$= \mathbf{h}^T (\text{Var}[\mathbf{X}_t \mathbf{b}_t] + 2\text{Cov}[\mathbf{X}_t \mathbf{b}_t, \mathbf{u}_t] + \text{Var}[\mathbf{u}_t]) \mathbf{h}. \quad (16)$$

Alphas and betas. We can rewrite the multi-factor model (13) in terms of the conventional alpha and beta parameters that arise in standard linear regression. More specifically, the alpha parameter $\alpha_t \in \mathbb{R}$ corresponds to the intercept of the linear model, the beta parameters $\beta_t \in \mathbb{R}^K$ corresponds to the K coefficients of the factor returns and $\eta_t \in \mathbb{R}$ is a zero-mean random variable:

$$R_t(\mathcal{I}, \mathcal{B}^{\text{free}}) = \alpha_t + \sum_{k=1}^K \beta_t^{(k)} b_t^{(k)} + \eta_t, \quad (17)$$

where $\alpha_t + \eta_t = \sum_{n=1}^N h^{(n)} u_t^{(n)}$ and $\beta_t^{(k)} = \sum_{n=1}^N h^{(n)} X_t^{(n,k)}$ for $k \in [K]$ and $t > 0$. Notably, it is not hard to see that the alpha and beta parameters of an investment $\mathcal{I} \in \mathbb{I}_N$ is just a weighted combination of the corresponding alpha and betas associated with the individual assets $\mathcal{A}_n \in \mathbb{A}_N$, respectively.

Structural assumptions. If we assume the following structural assumptions:

1. The factor loadings are deterministic.
2. The idiosyncratic risk of the assets \mathbf{u}_t are uncorrelated with the factor returns \mathbf{b}_t .
3. The idiosyncratic risk \mathbf{u}_t are uncorrelated with each other.

Then the variance of the excess returns reduces down to

$$\text{Var}[R_t(\mathcal{I}, \mathcal{B}^{\text{free}})] = \mathbf{h}^T (\mathbf{X}_t \text{Var}[\mathbf{b}_t] \mathbf{X}_t^T + \text{Var}[\mathbf{u}_t]) \mathbf{h}. \quad (18)$$

Equipped with this expression, one only needs to estimate $\mathcal{O}(K^2 + N)$ number of parameters in order to obtain the covariance matrix of interest $\Sigma_t = \mathbf{X}_t \text{Var}[\mathbf{b}_t] \mathbf{X}_t^T + \text{Var}[\mathbf{u}_t]$. Clearly,

this reduction in dimensionality is very desirable because the number of factors K is often several times smaller than the number of assets N . As a result, it is very common to see these structural assumptions being used together in practice.

Data processing. As with any data-driven model, some care has to be taken when setting up the exposure matrix. In particular, it is often advised to perform some level of standardisation, normalisation and outlier handling in order to construct the matrix. TODO:maybe ref something.

Choosing the factors. In general, the K factors in the linear model in (13) are chosen either manually or systematically based on a mixture of domain expertise and historical data analysis. There is no consensus on what the best factors to use are. In fact, there is a whole zoo of factors which have been suggested by different authors TODO:ref. Notably, the most famous factor models are the ones proposed by Fama and French [1993] etc. TODO:add more references.

Estimating the unknowns. The main parameters of the linear model (13) are the exposure matrix \mathbf{X}_t and the factor returns \mathbf{b}_t . In practice, both of these variables might not be known and therefore one might be interested in constructing an estimate of these unknowns based on what data is available. Depending on what we do know, the strategy that we take to estimate the corresponding unknowns would differ. For example, if we know \mathbf{X}_t , then we can estimate \mathbf{b}_t using cross-sectional regression. That is, we regress using data obtained at a specific slice in time. For instance, we can use a standard linear regression strategy or some (robust) variant thereof. Similarly, if we know \mathbf{b}_t , then we can estimate \mathbf{X}_t using time-series regression. That is, we regress using data obtained over many slices of time. In the case where both \mathbf{X}_t and \mathbf{b}_t are unknown, then we can try to estimate these two unknowns jointly using a statistical strategy such as principle component analysis.

3.1.2 Single factor model

When we have only a single factor, $K = 1$, then (13) reduces down to the single factor model

$$R_t(\mathcal{I}, \mathcal{B}^{\text{free}}) = \sum_{n=1}^N h^{(n)} X_t^{(n)} b_t + \sum_{n=1}^N h^{(n)} u_t^{(n)}, \quad (19)$$

for assets $n \in [N]$ and time $t > 0$.

CAPM. A special case of the single factor model is the Capital Asset Pricing Model (CAPM) where we make the following assumptions:

1. The factor return is equal to the excess market returns: $b_t = R_t(\mathcal{B}^{\text{market}}, \mathcal{B}^{\text{free}})$.
2. The expectation of the residuals are zero: $\mathbb{E}[u_t^{(n)}] = 0$ for $n \in [N]$.
3. The residuals are uncorrelated with each other: $\mathbb{E}[u_t^{(n)} u_t^{(n')}] = 0$ for $n \neq n' \in [N]$.

Following Grinold and Kahn [1999, Chapter 1], one can define the alpha and beta of a stock as the intercept and gradient of the least squares fit obtained by regressing the excess return of the investment with the excess market return:

$$R_t(\mathcal{I}, \mathcal{B}^{\text{free}}) = \alpha_t + \beta_t R_t(\mathcal{B}^{\text{market}}, \mathcal{B}^{\text{free}}) + \eta_t, \quad (20)$$

where η_t is the residual which is assumed to have zero mean and

$$\beta_t = \frac{\text{Cov}[R_t(\mathcal{I}, \mathcal{B}^{\text{free}}), R_t(\mathcal{B}^{\text{market}}, \mathcal{B}^{\text{free}})]}{\text{Var}[R_t(\mathcal{B}^{\text{market}}, \mathcal{B}^{\text{free}})]}. \quad (21)$$

Notice that the second CAPM assumption is in essence equal to suggesting that $\alpha_t = 0$. That is, there is no additional return gained by incurring more systematic risk than the return that we would expect from investing in the market. The only difference between investors is their risk tolerance. The amount of risk is dictated by the β_t parameter. When $\beta_t = 1$, we assume the same risk as the market; when beta is greater (or smaller) than one, we assume more (or less) risk than the market, respectively. Naturally, the more risk we take on, the greater the expected return we expect to get.

Weaknesses. A criticism about CAPM is that it is too simple in the sense that it is unlikely that the excess market return is the only factor which determines the systematic risk.

3.1.3 Arbitrage pricing theory

The Arbitrage Pricing Theory (APT) is a generalisation to CAPM. Instead of just using a single factor, **TODO: ross ref** proposes using the multi-factor model (13). Similarly with CAPM, the APT model assumes some additional properties on the expectations of the idiosyncratic risk and factors:

1. The expected values of the idiosyncratic risk are zero: $\mathbb{E}[u_t^{(n)}] = 0$ for all $n \in [N]$.
2. The residuals are uncorrelated with each other: $\mathbb{E}[u_t^{(n)} u_t^{(n')}] = 0$ for $n \neq n' \in [N]$.
3. The residuals are uncorrelated with the factor returns: $\mathbb{E}[u_t^{(n)} b_t^{(k)}] = 0$ for $n \in [N]$ and $k \in [K]$.

In contrast with CAPM, these assumptions are based on a well-known arbitrage argument. More specifically, suppose that the multi-factor model (13) holds and the APT assumptions do not. Then, whenever one manages to identify a mispriced investment (according to the factor model), one can potentially design a long-short type portfolio in order to generate returns without any expense—the precise strategy is outlined by [Grinold and Kahn \[1999, Chapter 7\]](#). In other words, these APT assumptions are sufficient enough to prevent these arbitrage opportunities from happening consistently.

Weaknesses. A criticism about the APT model is that it does not tell use what factors to use as opposed to the [Fama and French \[1993\]](#) models etc. **TODO:add ref.**

3.2 Machine learning

TODO: draft some preliminary ideas.

4 References

Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, February 1993. ISSN 0304-405X. doi: 10.1016/0304-405X(93)90023-5. URL <https://www.sciencedirect.com/science/article/pii/0304405X93900235>. Cited on pages 7 and 8.

Richard C. Grinold and Ronald N. Kahn. *Active Portfolio Management : A quantative approach for producing superior returns and selecting superior money managers*. McGraw Hill, New York, NY, 2nd edition edition, November 1999. ISBN 978-0-07-024882-3. Cited on pages 1, 2, 3, 4, 7, and 8.