On a Generalization of Markowitz Efficient Portfolios

Valentin Vankov Iliev

Sofia, 7 June 2016

Abstract

Given families of continuous objective functions $e = (e_p)_{p \in I}$ and continuous loss functions $v = (v_q)_{q \in J}$ on the set Δ_{n-1} of all portfolios of n financial assets, we define a preference relation $\succeq$ on Δ_{n-1} by the following rule: Any member of e is an $\succeq$-increasing and any member of v is an $\succeq$-decreasing function. Under the condition that there are no short sales, any portfolio $x \in \Delta_{n-1}$ is $\succeq$-dominated by a $\succeq$-maximal portfolio $y \in \Delta_{n-1}$: $y \succeq x$. Thus, we obtain a generalization of Harry Markovitz's classical mean-variance portfolio theory. The applications of this generalized theory are based on the main shortcoming of the classical approach: For the most part the distributions of returns are non-normal so can not be captured purely by mean and variance. The generalization allows to adapt a large number of families of objective/loss functions to the appropriate distributions of returns and to design these distributions.

In this presentation a *financial market* consists of:

- *Assets*: $1, 2, \ldots, n$.
- *Return* r_i on asset i for a given period of time — this is a random variable, which, without any reinvestment, is the *rate of return*: profit on an investment over a period of time, expressed as a proportion of the original investment.
- *Expected return* on asset i for a given period of time: $\mu_i = \mathrm{E}(r_i)$, the mean of the random variable r_i.
- *Risk*, or, *volatility*, or, *variance* of asset i for a given period of time: $\sigma_i^2 = \mathrm{Var}(r_i)$, the variance of the random variable r_i.

From now on we assume (after eventual rearranging of the assets) that $0 < \sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_n$ and that no one of the random variables r_i is a linear combination of the others.

Introductory Remarks: Portfolios

A *n-asset portfolio* is a sequence $x = (x_1, x_2, \dots, x_n)$ of non-negative numbers with $x_1 + x_2 + \dots + x_n = 1$ (no $x_i < 0$, that is, no "short sales"). Here x_i is the relative amount of money invested in asset i. We denote by Δ_{n-1} the set of all n-asset portfolios.

- (a) *Return* $r(x)$ on the portfolio x for a given period of time: $r(x) = x_1 r_1 + x_2 r_2 + \dots + x_n r_n$.
- (b) *Expected return* $e(x)$ on the portfolio x for a given period of time: $e(x) = \mathrm{E}(r(x)) = x_1 \mu_1 + x_2 \mu_2 + \dots + x_n \mu_n$.
- (c) *Risk*, or, *volatility*, or, *variance* $v(x)$ of the portfolio x for a given period of time:

$$v(x) = \mathrm{Var}(r(x)) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij},$$

where $\sigma_{ij} = \mathrm{Cov}(r_i, r_j)$ is the covariance of r_i and r_j.

Example

Let us suppose that 30% of the portfolio $x = (x_1, x_2, x_3)$ are invested in Ford Motor Co ($x_1 = 0.30$), 45% of the portfolio are invested in Facebook Inc ($x_2 = 0.45$), and 25% of the portfolio are invested in Apple Inc ($x_3 = 0.25$). Note that there are not reinvestments.

Let us also suppose that the return on Ford Motor Co is $r_1 = 0.06$, the return on Facebook Inc is $r_2 = 0.12$, and the return on Apple Inc is $r_3 = 0.08$ for this time period.

Then the return $r(x)$ of the portfolio x is

$$r(x) = x_1 r_1 + x_2 r_2 + x_3 r_3 =$$

$$(0.30)(0.06) + (0.45)(0.12) + (0.25)(0.08) = 0.092$$

for the same time period.

In his fundamental 1952 paper [1], Harry Markowitz stated several theses:

- 1) "The portfolio with maximum expected return is not necessarily the one with minimum variance".
- 2) "There is a rate at which the investor can gain the expected return by taking on variance, or reduce variance by giving up expected return".
- 3) "The returns from securities are too intercorrelated. Diversification can not eliminate all variance".
- 4) "*Expected returns–variance of returns (e-v) rule*: The investor would (or should) want to select one of those portfolios with minimum v for given e or more and maximum e for given v or less".

In [1] Markowitz presents geometrically the existence of the portfolios from (e-v) rule (so called *efficient portfolios*) in the case of 3 and 4 assets. Later, in his monograph [2, Part III, Ch. VII, VIII, Appendix A], he describes again meticulously 3- and 4-asset case and gives algorithms (again omitting proofs) for finding the efficient portfolios in the general case of n assets. In 1990 Harry Markowitz received (together with Merton Miller and William Sharpe) Alfred Nobel Memorial Prize in Economic Sciences for his works on efficient portfolios.

In accord with (e-v) rule, the portfolio $x \in \Delta_{n-1}$ is efficient when

$$e(x) = \max_{y \in \Delta_{n-1}, v(y) \leq v(x)} e(y) \text{ and } v(x) = \min_{y \in \Delta_{n-1}, e(y) \geq e(x)} v(y). \quad (1)$$

In other words, for any portfolio $y \in \Delta_{n-1}$ the inequality $v(y) \leq v(x)$ implies the inequality $e(x) \geq e(y)$ and the inequality $e(y) \geq e(x)$ implies the inequality $v(x) \leq v(y)$. The negation of the last statement is: There exists $y \in \Delta_{n-1}$ such that

$$[(e(x) < e(y)) \text{ and } (v(x) \geq v(y))] \text{ or}$$

$$[(e(x) \leq e(y)) \text{ and } (v(x) > v(y))]. \quad (2)$$

The formula (2) defines a binary relation $y \succ x$, read "y is definitely better than x". Thus, there exists portfolio y with $y \succ x$, that is, definitely better than x.

Therefore the formula (1) means that there exists no portfolio y with $y \succ x$, that is, the portfolio x is (one of the) best, or, using mathematical language, it is *maximal*. Using Markowitz's terminology, the maximal portfolios x are said to be *efficient*. It is natural to say that the return e (respectively, the risk v) *improves* when e strictly increases (respectively, v strictly decreases). On the other hand, we say that the return e (respectively, the risk v) *gets worse* when e strictly decreases (respectively, v strictly increases). Thus, a portfolio is efficient if we can not improve one of these two ingredients without getting worse the other.

Taking only the return and the risk of a portfolio into account, it is natural not to distinguish portfolios x and y for which $e(x) = e(y)$ and $v(x) = v(y)$; in this case we write $x \simeq y$ (an equivalence relation). We also write $y \succeq x$, read "y is better than x" when $y \succ x$ or $x \simeq y$. We call the relation $\succeq$ *Markowitz's preference relation*. Thus, the portfolio x is efficient (that is, maximal) if for any portfolio y the relation $y \succeq x$ implies the relation $y \simeq x$. After identifying the equivalent portfolios, we obtain a compact region in Markowitz's $e - v$ space. In his paper [1, p. 87] Markowitz states that the set of all efficient portfolios in $e - v$ space is "a series of connected line segments" and presents an algorithm for moving along these line segments. In order to illustrate his ideas, below we consider the simplest (but not trivial) case of 2-asset portfolios.

We have $x = (x_1, x_2)$, $x_1 \geq 0$, $x_2 \geq 0$, and $x_1 + x_2 = 1$. Let us substitute $x_1 = t$, $0 \leq t \leq 1$. Hence $x_2 = 1 - t$, the return $r(x)$ of the portfolio x is $r(x) = tr_1 + (1 - t)r_2$, and its numerical characteristics (as functions in t) are:

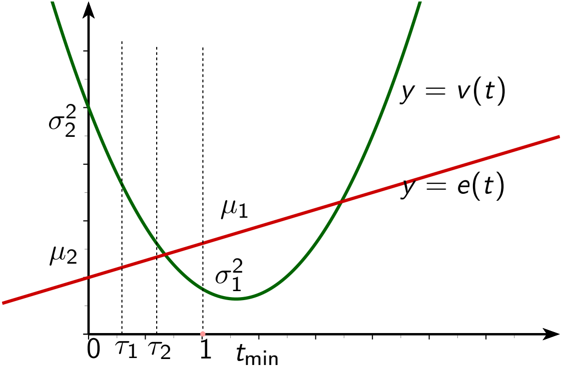
- Expected return $e(t) = (\mu_1 - \mu_2)t + \mu_2$— a linear function.
- Risk $v(t) = (\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})t^2 + 2(\sigma_{12} - \sigma_2^2)t + \sigma_2^2$ — a quadratic function with negative discriminant.

In this case the definition (1) of an efficient portfolio $x(\tau)$ has the form

$$e(\tau) = \max_{0 \leq t \leq 1, v(t) \leq v(\tau)} e(t) \text{ and } v(\tau) = \min\nolimits_{0 \leq t \leq 1, e(t) \geq e(\tau)} v(t). \quad (3)$$

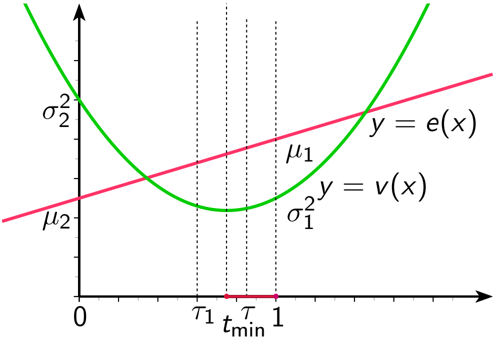
Since the graph of the return $e(t)$ is a straight line and the graph of the risk $v(t)$ is a parabola, using our knowledge from high school, we can easily analyze all possible cases.

Two-Security Portfolios: Figure (1) $\sigma_1^2 \leq \sigma_{12}$ and $\mu_1 \geq \mu_2$



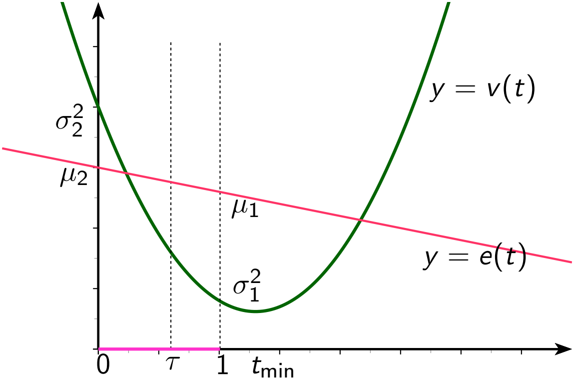
$x(\tau_2) \succ x(\tau_1) \Longrightarrow x(1)$ is the only efficient portfolio.

Two-Security Portfolios: Figure (2) $\sigma_1^2 > \sigma_{12}$ and $\mu_1 \geq \mu_2$



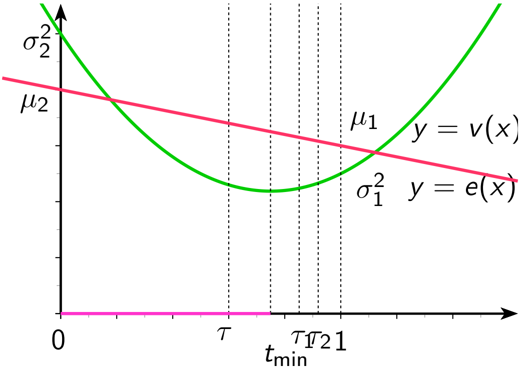
$x(t_{\min}) \succ x(\tau_1)$ for $\tau_1 \in [0, t_{\min})$ and $e(\tau) = \max_{t \leq \tau} e(t)$, $v(\tau) = \min_{\tau \leq t} v(t) \implies x(\tau)$, $\tau \in [t_{\min}, 1]$, is the set of efficient portfolios.

Two-Security Portfolios: Figure (3) $\sigma_1^2 \leq \sigma_{12}$ and $\mu_1 < \mu_2$



$e(\tau) = \max_{\tau \leq t} e(t)$ and $v(\tau) = \min_{t \leq \tau} v(t) \Longrightarrow x(\tau)$ is efficient portfolio for all $\tau \in [0, 1]$.

Two-Security Portfolios: Figure (4) $\sigma_1^2 > \sigma_{12}$ and $\mu_1 < \mu_2$



$e(\tau) = \max_{\tau \le t} e(t)$ and $v(\tau) = \min_{t \le \tau} v(t) \implies x(\tau)$ is efficient portfolio for all $\tau \in [0, t_{\min}]$.

$x(\tau_1) \succ x(\tau_2)$ for all $\tau_1 \in [t_{\min}, \tau_2) \implies x(\tau_2)$ is not an efficient portfolio for all $\tau_2 \in (t_{\min}, 1]$.

The above considerations confirm Markowitz's statement that the set of efficient portfolios (that is, portfolios which are maximal with respect to Markowitz's preference relation $\succeq$) is a series of connected line segments and show, in particular, that this set is not empty.

Question

Keeping Markowitz's idea of efficiency, is it possible to increase the number of functions of type e and the number of functions of type v, simultaneously maximizing e's and minimizing v's?

The answer is "yes", it is possible, and below we explain how to do that.

Let $e = (e_p)_{p \in I}$ and $v = (v_q)_{q \in J}$ be two families of continuous real functions on the set of all portfolios Δ_{n-1}. Let us suppose that Δ_{n-1} is furnished with a preference relation $\succeq$ produced by the families e and v in the following way:

- We write $y \succeq x$ if $e_p(x) \leq e_p(y)$ and and $v_q(x) \geq v_q(y)$ for all $p \in I$, $q \in J$.
- For the symmetric part $\simeq$ of the preference relation $\succeq$ one has $y \simeq x$ if $e_p(x) = e_p(y)$ and $v_q(x) = v_q(y)$.
- For the anti-symmetric part $\succ$ of the preference relation $\succeq$ one has $y \succ x$ if $y \succeq x$ and not $y \simeq x$. This means $y \succeq x$ and either there exists index $p_0 \in I$ with $e_{p_0}(x) < e_{p_0}(y)$ or there exists index $q_0 \in J$ with $v_{q_0}(x) > v_{q_0}(y)$.
- A portfolio x is called $\succeq$-*maximal* if there is no portfolio y with $y \succ x$.

Theorem

For any portfolio $x \in \Delta_{n-1}$ there exists an $\succeq$-maximal element $y \in \Delta_{n-1}$ with $y \succeq x$.

With a view to explain the mechanism of proving the existence of a $\succeq$-maximal portfolio, let us remind the *Principle of Mathematical Induction* from high school. Let $\mathcal{P}(n)$ be a statement which depends on a natural number $n \in \{1, 2, 3, \dots\}$. If $\mathcal{P}(1)$ is true and if the validity of $\mathcal{P}(m)$ implies the validity of $\mathcal{P}(m+1)$, then the statement $\mathcal{P}(n)$ is true for all natural numbers n. Equivalently: If $\mathbb{N}$ is the set of natural numbers, then every non-empty subset $A \subset \mathbb{N}$ has a minimal element $a \in A$. The *Axiom of Choice* says that given a set X, for any non-empty subset $A \subset X$ one can fix an element $a \in A$. In this way we can use the principle of mathematical induction (appropriately generalized) in any set X in order to prove the above theorem.

The set Δ_{n-1} of all portfolios is, in fact, the $(n-1)$-dimensional simplex which is a compact set in the n-dimensional real vector space. It is well known, that Δ_{n-1} is a quasi-compact and sequentially compact topological space.

Now, let X be a quasi-compact and sequentially compact topological space and let $e = (e_p)_{p \in I}$, $v = (v_q)_{q \in J}$ be two families of continuous real functions on X. The families e and v produce a preference relation $\succeq$ on X via the same rule: $y \succeq x$ if $e_p(x) \leq e_p(y)$ and and $v_q(x) \geq v_q(y)$ for all $p \in I$, $q \in J$. The previous Theorem is a particular case of the following much more general statement.

Theorem

For any element $x \in X$ there exists an $\succeq$-maximal element $y \in X$ with $y \succeq x$.

Example

Let B_{n-1} be an $(n-1)$-dimensional closed ball in the hyperplane $x_1 + x_2 + \cdots + x_n = 1$ of $\mathbb{R}^n$ with $\Delta_{n-1} \subset B_{n-1}$. The members of B_{n-1} can be considered as portfolios with bounded short sales. The ball B_{n-1} is a a quasi-compact and sequentially compact topological space, so the generalized Theorem yields that for any portfolio $x \in B_{n-1}$ there exists an $\succeq$-maximal element $y \in B_{n-1}$ with $y \succeq x$. In other words, any portfolio is $\succeq$-dominated by an efficient portfolio.

Below, we remind some notions from statistics and give examples of application of the above Theorem.
Given the integer $\ell \geq 2$, the *ℓ-th central moment* of the random variable $r(x)$ is $E((r(x) - E(r(x)))^{\ell})$. The standard variance is the second central moment $v(x) = E((r(x) - E(r(x)))^2)$ of $r(x)$ and it is a quadratic form in $x_1, \ldots, x_n$. The third central moment $E((r(x) - E(r(x)))^3)$ is a cubic form and the fourth central moment $E((r(x) - E(r(x)))^4)$ is a form of degree 4 in $x_1, \ldots, x_n$.

Given $x \in \Delta_{n-1}$ and $t \in \mathbb{R}$, we set

$$F_x(t) = P(\{m \in S \mid r(x)(m) < t\}).$$

The function $F_x \colon \mathbb{R} \to [0,1]$ is said to be the *cumulative distribution function* of the random variable $r(x)$. We assume that $r(x)$ is a continuous random variable with continuous *density function* $f_x(t)$, so $F_x(t) = \int_{-\infty}^{t} f_x(\tau) d\tau$ and $F'_x(t) = f_x(t)$. In particular, the functions $F_x(t)$ are continuous in t. If, in addition, the random variables $r_1, \ldots, r_n$ are independent (in particular, not intercorrelated), then the distribution functions $F_x(t)$ are continuous also with respect to x.

The portfolio $x \in \Delta_{n-1}$ is said to be *first order stochastically dominated* by portfolio $y \in \Delta_{n-1}$ if $F_y(t) \leq F_x(t)$ for all $t \in \mathbb{R}$.

Statistics preliminaries: Skewness and Kurtosis

We set

$$\mathrm{Skew}(r(x)) = \frac{E((r(x) - E(r(x)))^3)}{\mathrm{Var}(r(x))^{\frac{3}{2}}}$$

to be the *skewness* and

$$\mathrm{Kurt}(r(x)) = \frac{E((r(x) - E(r(x)))^4)}{\mathrm{Var}(r(x))^2} - 3$$

to be the *kurtosis*, or, *excess kurtosis* of the random variable $r(x)$.
If the random variable $r(x)$ is normal, then

$$\mathrm{Skew}(r(x)) = \mathrm{Kurt}(r(x)) = 0.$$

Example

In case $I = \{1\}$, $J = \emptyset$, the function $e = e_1$ can be considered as an utility function on Δ_{n-1} and $\succeq$ is the corresponding preference relation with negatively transitive asymmetric part $\succ$.

Example

In case

$$e(x) = E(r(x)) = x_1\mu_1 + \cdots + x_n\mu_n,$$

$$v(x) = \mathsf{Var}(r(x)),$$

we obtain the classical Markowitz's setup.

Example

In case

$$e(x) = e_0 = const,$$

$$v(x) = \mathrm{Var}(r(x)),$$

the generalized efficient portfolios are exactly those x for which the risk attains its *absolute minimum*:

$$v(x) = \min_y v(y).$$

Example

In case

$$e(x) = E(r(x)) = x_1\mu_1 + \cdots + x_n\mu_n,$$

$$v(x) = v_0 = const,$$

the generalized efficient portfolios are exactly those x for which the return attains its *absolute maximum*:

$$e(x) = \max_y e(y).$$

Example

In case

$$e_1(x) = E(r(x)),$$

$$v_1(x) = \mathsf{Var}(r(x)),\ \ v_2(x) = \mathsf{Skew}^2(r(x)),$$

we simultaneously maximize the expected return $E(r(x))$ and minimize the volatility $\mathsf{Var}(r(x))$ and the square of the skewness $\mathsf{Skew}(r(x))$ of the return $r(x)$, thus balancing the tails of its distribution.

Example

In case

$$e_1(x) = E(r(x)),$$

$$v_1(x) = \mathrm{Var}(r(x)),\ \ v_2(x) = \mathrm{Kurt}^2(r(x)),$$

we simultaneously maximize the expected return $E(r(x))$ and minimize the volatility $\mathrm{Var}(r(x))$ and the square of the kurtosis $\mathrm{Kurt}(r(x))$ of the return $r(x)$.

Example

In case

$$e_1(x) = E(r(x)),$$

$$v_1(x) = \mathsf{Var}(r(x)),\ v_2(x) = \mathsf{Skew}^2(r(x)),\ v_3(x) = \mathsf{Kurt}^2(r(x)),$$

we simultaneously maximize the expected return $E(r(x))$ and minimize the volatility $\mathsf{Var}(r(x))$, the square of skewness $\mathsf{Skew}(r(x))$, and the square of kurtosis $\mathsf{Kurt}(r(x))$ of the return $r(x)$, thus mimicking the normal distribution.

Examples

Example

In case

$$e_1(x) = E(r(x)),\ \ e_2(x) = \mathsf{Skew}(r(x))$$

$$v_1(x) = \mathsf{Var}(r(x)),\ \ v_2(x) = \mathsf{Kurt}(r(x)),$$

we simultaneously maximize the expected return $E(r(x))$ and the skewness $\mathsf{Skew}(r(x))$, and minimize the volatility $\mathsf{Var}(r(x))$ and the kurtosis $\mathsf{Kurt}(r(x))$, thus moving more probability mass to the right tail of the distribution.

Example

Let $\mathsf{Sh}(r(x)) = \frac{\mathsf{E}(r(x))}{\sqrt{\mathsf{Var}(r(x))}}$ be the Sharpe ratio. In case

$$e_1(x) = \mathsf{Sh}(r(x)), \;\; e_2(x) = \mathsf{Skew}(r(x)),$$

$$v_1(x) = \mathsf{Kurt}(r(x)),$$

we simultaneously maximize the Sharpe ratio $\mathsf{Sh}(r(x))$ and the skewness $\mathsf{Skew}(r(x))$, and minimize the kurtosis $\mathsf{Kurt}(r(x))$.

Example

Let the random variables $r_1, \ldots, r_n$ be independent. In case

$$e(x) = E(r(x)),$$

$$v(x) = \mathrm{Var}(r(x)),\ v_t(x) = F_x(t),\ t \in \mathbb{R},$$

we simultaneously maximize the expected return $E(r(x))$ and the first-order stochastic dominance and minimize the volatility $\mathrm{Var}(r(x))$.

H. Markowitz, Portfolio Selection, The Jornal of Finance, vol. 7, No 1, (1952), 77 - 91.

H. Markowitz, *Portfolio Selection, 2nd ed.*, Blackwell 1991.

S. Roman, *Introduction to the Mathematics of Finance: From Risk Management to Options Pricing*, Springer-Verlag, 2004.

Thank you!