

Evolutionary Stochastic Portfolio Optimization and Probabilistic Constraints

Ronald Hochreiter

Department of Finance, Accounting, and Statistics, WU Vienna University of Economics and Business

CompStat 2010. Paris, France. August 2010.

From Deterministic to Stochastic Programming

To calculate an optimal decision $x \in \mathbb{R}^n$ in a **deterministic** world solve

$$\begin{aligned} & \text{minimize } x : f(x) \\ & \text{subject to } x \in \mathcal{X}, \end{aligned}$$

with some deterministic cost function $f(\cdot)$. In a **stochastic** world replace deterministic parameters by **probability distributions or stochastic processes** and optimize over a **probability functional** \mathbb{F} , i.e.

$$\begin{aligned} & \text{minimize } x : \mathbb{F}(f(x, \Xi)) \\ & \text{subject to } (x, \Xi) \in \mathcal{X}, \end{aligned}$$

on some probability space Ξ , where $f(\cdot, \cdot)$ is a stochastic cost function.

The probability functional \mathbb{F} is commonly the expectation, some **(coherent) risk measure**, or a weighted combination: link to Mathematical Finance.

Stochastic Programming - Discrete Uncertainty

Analytical solutions of stochastic programs using continuous distribution functions may only be computed using **unrealistic assumptions and simplifications** of the optimization model as well as the **uncertainty model**.

To build **real-world decision support applications**, uncertainty has to be expressed in some **discrete representation** to obtain numerical solutions, which depends on the stochastic stage structure (static vs. dynamic):

- Single-stage, two-stage: (multi-variate) **discrete** probability distribution (**scenario set**).
- Multi-stage: **discrete-time, discrete-space** (multi-variate) stochastic process. Approximation of non-anticipativity constraints (**scenario tree**).

Stochastic Programming - Integrating Uncertainty

- **Optimization model** extended to specify the event space at each decision stage.
 - integrate **objective** real-world constraints and dynamics, i.e. the event handling (**quantitative** aspects of the solution).
- **Uncertainty model** commonly chosen independently from optimization model.
 - reflect **subjective** beliefs of the decision taker (**qualitative** aspects of the solution).
 - due to the subjective nature a wide range of heuristic methods and procedures exist, mainly sampling-based or clustering-based.

Interdisciplinary Approach

Motivation: Generalization of financial portfolio optimization using contemporary risk management techniques by combining different areas of research:

- Financial Engineering - Portfolio selection [Markowitz],
- Mathematical Finance - (Coherent) Risk measures [Artzner et al.],
- Operations Research - **Stochastic Optimization**.

One of the simplest Stochastic Programming examples: static single-stage (no recourse action), but powerful in terms of modeling opportunities and applications.

Financial engineering - Portfolio selection

- Classical bi-criteria portfolio optimization problem (Markowitz 1952, 1958).
- Calculate an optimal investment portfolio out of set of assets (or asset categories) \mathcal{A} with finite cardinality $a = |\mathcal{A}|$.
- Given a co-variance matrix \mathbb{C} and a return vector M for all assets $\in \mathcal{A}$ find the optimal portfolio x by solving the quadratic optimization problem:

$$\begin{aligned} \text{minimize } x : & \quad x \mathbb{C} x^T \\ \text{subject to } & \quad x \times M \geq \mu \\ & \quad x \in \mathcal{X}. \end{aligned}$$

Let μ be the minimum return required by the investor. M and \mathbb{C} are calculated from historical data and/or generated using expert opinion.

- **Variance** is defined as the **main risk factor** and is **minimized**.

Operations Research - Stochastic Portfolio Optimization

Instead of using a vector of asset means M and a co-variance matrix \mathbb{C} , a set of asset return scenarios \mathcal{S} is used, i.e. a multi-variate probability distribution.

- \mathcal{S} has finite cardinality $s = |\mathcal{S}|$.
- Each scenario $s_i \in \mathcal{S}$ is equipped with a weight / probability $p_i \geq 0$, with $\sum_{j=1}^s p_j = 1$.
- \mathcal{S} is the result of scenario generation: **Sampling historical data and/or generated using expert opinion** (time series model estimation and simulation, sampling), and using an (optimal) **approximation scheme**.

Better dependency modeling instead of using a co-variance matrix, subtle weighting of scenarios possible (using some indices like the VIX).

Stochastic Portfolio Optimization

Let $x \in \mathbb{R}^a$ be some portfolio, with budget normalization, i.e. $\sum_{a \in \mathcal{A}} x_a \leq 1$.

Rewrite \mathcal{S} as matrix S : Loss distribution ℓ for some portfolio x is $\ell_x = \langle x, S \rangle$.

Two dimensions: Value (Return) $\mathbb{E}(\ell_x)$ and Risk $\mathbb{F}^\rho(\ell_x)$

Optimization problem:

$$\begin{aligned} & \text{maximize } x : \mathbb{E}(\ell_x) \\ & \text{minimize } x : \mathbb{F}^\rho(\ell_x) \\ & \text{subject to } x \in \mathcal{X} \end{aligned}$$

Multi-criteria reformulation with risk-aversion parameter κ

$$\text{maximize } x : \mathbb{E}(\ell_x) - \kappa \mathbb{F}^\rho(\ell_x), \text{ subject to } x \in \mathcal{X}$$

Stochastic Portfolio Optimization - Constraints

Most common constraints, included into \mathcal{X} are

- upper/lower bounds: $x_a \geq l_a, x_a \leq u_a \forall a \in \mathcal{A}$
(e.g. no short selling allowed: $l_a = 0$, i.e. $x_a \geq 0, \forall a \in \mathcal{A}$),
- minimum expected profit: $\mathbb{E}(\ell_x) \geq m$,
- maximum expected risk: $\mathbb{F}^\rho(\ell_x) \leq r$,
- cardinality constraints: $\#(x_a > 0) \leq K, K \in \mathbb{N}$,
- transaction costs,
- ...

Stochastic Portfolio Optimization - Risk Measures & Generalization

Commonly applied risk measures \mathbb{F}^ρ for portfolio management:

- Standard Deviation (resembles Markowitz): $\mathbb{F}^\rho = \sigma = \sqrt{\sum_{i \in \mathcal{S}} p_i (\ell_i - \mathbb{E}(\ell))^2}$.
- Value at Risk (VaR) at level $(1-\alpha)$ is the α -Quantile of loss distribution (important for regulatory purposes, easy to evaluate, hard to optimize):

$$\mathbb{F}^\rho = \text{VaR}_\alpha = \inf\{l \in \mathbb{R} : \mathbb{P}(l > \ell) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R} : F_l(\ell) \leq \alpha\}$$

- Conditional Value at Risk (CVaR), Expected Shortfall (substitute for VaR - linear programming reformulation, coherent risk measure): $\mathbb{F}^\rho = \text{CVaR}_\alpha = \mathbb{E}(\ell | \ell \leq \text{VaR}_\alpha)$

Different reformulations are available (in the above cases: quadratic, d.c., linear). Omega.

Evolutionary Algorithm

We consider the single objective version of the optimization problem (minimize risk subject to a minimal required return)

$$\begin{aligned} & \text{minimize} && \mathbb{F}(\ell_x) \\ & \text{subject to} && \mathbb{E}(\pi_x) > \mu \end{aligned} \tag{1}$$

Standard portfolio constraints (budget normalization & upper and lower bounds)

$$\begin{aligned} & \text{subject to} && \sum_{a \in \mathcal{A}} x_a = 1 \\ & && l \leq x_a \leq u \quad \forall a \in \mathcal{A} \end{aligned} \tag{2}$$

Optimization can be tricky (real-world constraints), Evaluation is cheap!

Genotype encoding

Each gene consists of of two parts:

1. One determines the amount of budget to be distributed to each selected asset, and
2. the other determines in which assets to invest.

The first part g_1 consists of a predefined number b of real values between 0 and 1 and the second part g_2 is encoded as a bit-string of the size of the amount of assets.

Example with $b = 6$ and $a = 3$: (0.31, 0.62, 0.39, 0.54, 0.22, 0.04, 1, 0, 1)

Invest into asset 1 and 3 buckets (1,2,1,2,1,1), i.e. each 0.5 of the budget.

Genotype encoding

Example. Let us define a bucket size of $b = 10$ and consider that we want to select the optimal portfolio out of $a = 5$ assets. A random chromosome with a fixed number of 3 asset picks may then consist of the following two parts g_1 and g_2 .

$$\begin{aligned}g_1 &= (0.4893, 0.3377, 0.9001, 0.3692, 0.1112, 0.7803, 0.3897, 0.2417, 0.4039, 0.0965), \\g_2 &= (1, 1, 0, 0, 1).\end{aligned}$$

If we re-map these two parts, we obtain the buckets $(2, 2, 3, 2, 1, 3, 2, 1, 2)$ and receive the following portfolio

$$x = (0.3, 0.5, 0, 0, 0.2),$$

which is a valid portfolio composition, and can be used to calculate the loss function and to conduct a full evaluation of the respective portfolio optimization formulation within the evolutionary optimization process.

Evolutionary operators

- Elitist selection. A certain number o_1 of the best chromosomes will be used within the next population.
- Crossover. A number o_2 of crossovers will be added, both 1-point crossovers (g_1 and g_2) and intermediate crossovers (only g_1).
- Mutation. o_3 mutated chromosomes will be added.
- Random addition. Furthermore o_4 randomly sampled chromosomes are added, which are also used for creating the initial population.

The specific number of operators $o = (o_1, o_2, o_3, o_4)$ has to be determined for each concrete number of assets a as well as the parameter b .

Example: Probabilistic constraints

Let the risk be the variance (resemble classic Markowitz approach): $\mathbb{F}(\ell_x) = \text{Variance}(\ell_x)$

Optimization is easy (quad.prog.), but what if we add a probabilistic constraint?

$$\text{subject to } \mathbb{P}(\pi_x \leq \delta) \leq \varepsilon.$$

Let the fitness value which we aim to minimizing be f . We calculate the penalty p by

$$p = f \times (\mathbb{P}(\pi_x \leq \delta) - \varepsilon) \times \gamma,$$

where γ is a factor to control the penalization level. The fitness value used for evolutionary optimization purposes is thus given by $f' = f + p$. Such a constraint can be implemented and handled conveniently.

Numerical results

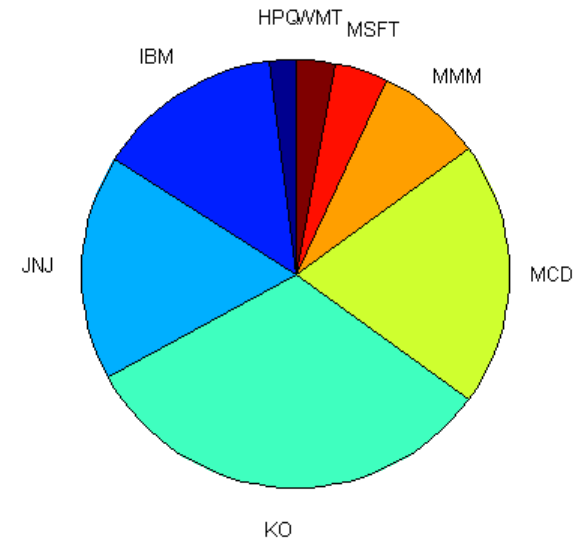
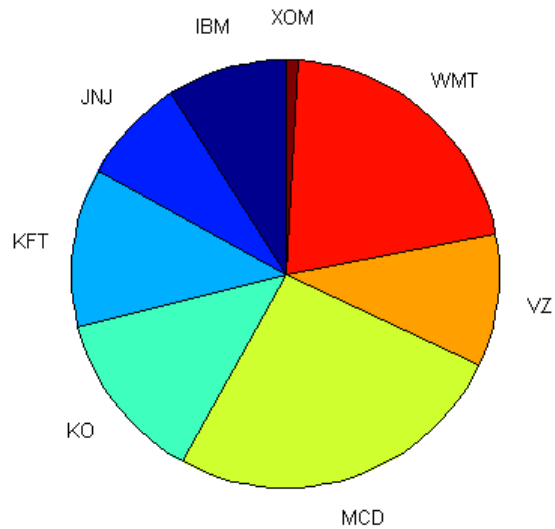
Implemented using MatLab 2008b without using any further toolboxes.

30 stocks of the Dow Jones Industrial Average at the beginning of 2010, i.e. the ticker symbols AA, AXP, BA, BAC, CAT, CSCO, CVX, DD, DIS, GE, HD, HPQ, IBM, INTC, JNJ, JPM, KFT, KO, MCD, MMM, MRK, MSFT, PFE, PG, T, TRV, UTX, VZ, WMT, XOM. Daily data from each trading day in 2009 has been used. Weekly returns have been calculated.

$b = 100$ buckets which are distributed to the respective asset picks, such that each chromosome has a length of $b + a = 130$. The initial population consists of 1000 random chromosomes. The operator structure $o = (100, 420, 210, 100)$.

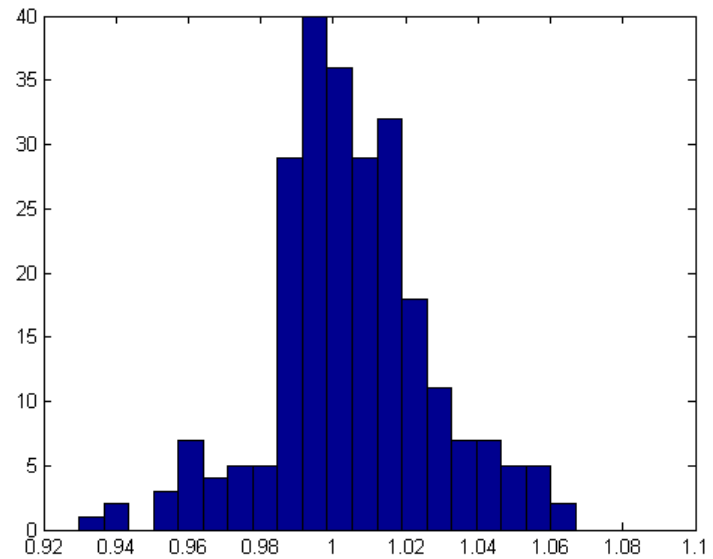
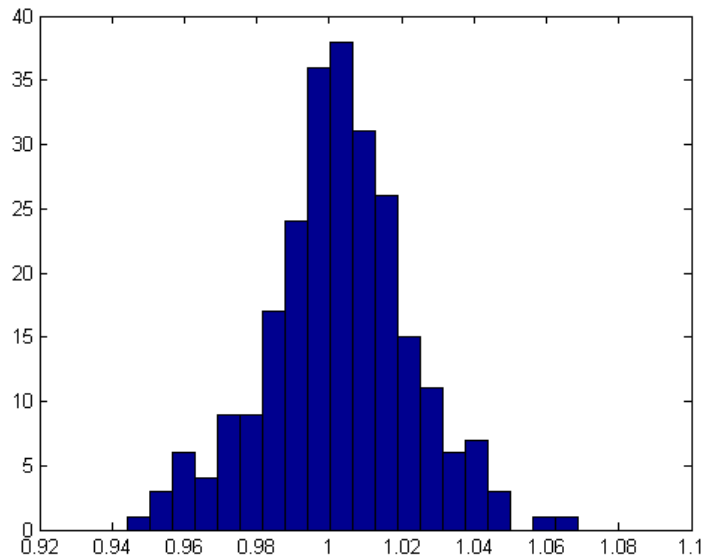
First, we optimize without using the probabilistic constraints, i.e. the main optimization problem using $\mu = 0.001$. Then we add the probabilistic constraint with $\delta =$ and $\varepsilon = 0.1$. These values have been chosen after analyzing the resulting loss distribution.

Numerical results - DJIA ($a = 30$)



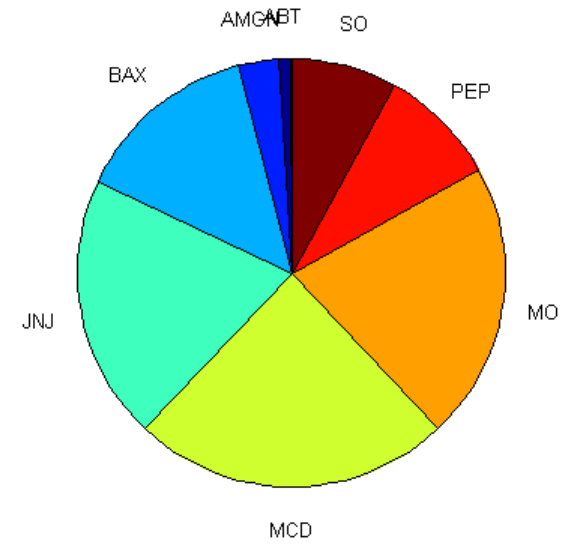
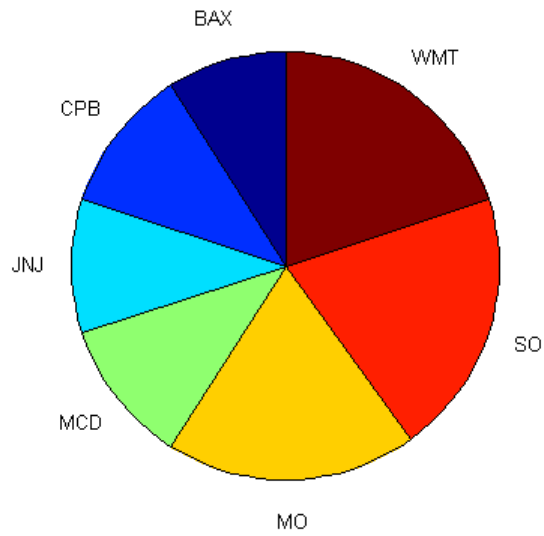
Diversification grade is not changed, KFT, VZ, XOM are dropped and HPQ, MSFT, MMM are picked.

Numerical results - DJIA ($a = 30$)

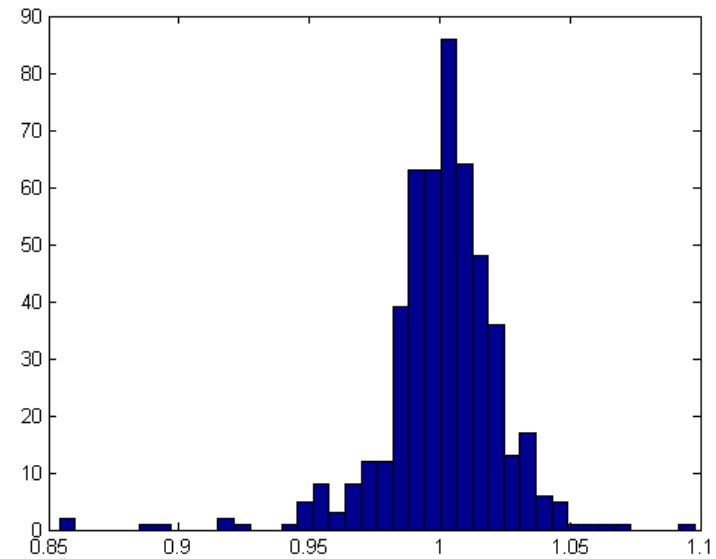
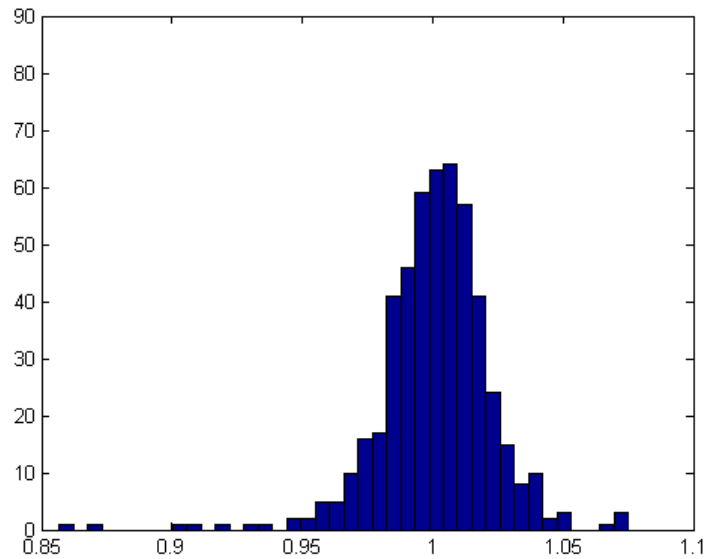


	P_1 (no \mathbb{P})	P_2	$P_3(1/N)$
Mean	0.0024	0.0051	0.0062
Std.Dev.	0.02	0.0225	0.0398
Prob.	0.1774	0.1089	0.2702

Numerical results - S&P100 ($a = 100$)



Numerical results - S&P100 ($a = 100$)



	P_1 (no \mathbb{P})	P_2	$P_3(1/N)$
Mean	0.0011	0.0009	0.0007
Std.Dev.	0.0220	0.0231	0.0451
Prob.	0.1677	0.1238	0.3134

Contact - Ronald Hochreiter

Snail Department of Finance, Accounting and Statistics
Institute for Statistics and Mathematics
WU Vienna University of Economics and Business
Augasse 2-6, A-1090 Vienna, Austria

Web *<http://www.hochreiter.net/ronald/research/>*

Email ronald.hochreiter@wu.ac.at

MatLab code is available on request!