Working Paper

Stochastic Optimization of a Multi-currency Bond Portfolio

Tuula Huoponen

WP-94-98
August 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 71521 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Stochastic Optimization of a Multi-currency Bond Portfolio

Tuula Huoponen

WP-94-98
August 1994

Working Papers are interim reports on work of the International Institute for Applied Systems Analysis and have received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute or of its National Member Organizations.

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 71521 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Abstract

This paper presents a stochastic optimization approach for the management of multi-currency government bond portfolio. This practical problem of optimal fund allocation is formulated as a linearly constrained two-stage model where the parameter values are not known with certainty but depend on future course of underlying stochastic economic variables. The model differs from the standard two-stage formulation as data for the second stage problem is uncertain as well. The objective is to maximize the expected utility of the market value of the portfolio at the end of the planning horizon. To solve the problem, we employ a stochastic quasigradient method by Ermoliev. For the optimality test, upper and lower estimates for the optimal objective function value are developed based on a given confidence level. According to initial numerical results, the convergence to a satisfactory near-optimal solution is considered sufficiently fast for a practical application.

Key words: Stochastic programming, interest rate risk, exchange rate risk.

Acknowledgements

I am very grateful to Markku Kallio for his excellent research support during the YSSP 1994. The implementation of the test application in such a short time also benefitted from the helpful guidance of Yuri Ermoliev and Alexei Gaivoronski. Additionally, I would like to thank A.P. Wierzbicki for his useful comments and Andrzej Ruszczyński for providing the random number generator for the computer application.

Stochastic Optimization of a Multi-currency Bond Portfolio

Tuula Huoponen

1. Introduction

This paper concentrates on decision making pertaining to the optimal management of a multi-currency bond portfolio. Practical situations where such decision problems arise are eg. investing the foreign currency reserves of a central bank, deciding on an optimal mix of currencies for project financing, and determining an optimal allocation for an investment fund to invest the shareholders' funds in bonds denominated in different currencies.

In this paper, we consider the combined effect of exchange rate and interest rate uncertainty on managing a portfolio that consists of domestic and foreign government bonds. Government bonds can be broadly defined as contractual obligations of governments to repay the borrowed funds with interest. By restricting the investments only in government bonds we assume we can omit the credit risk. Thus, the financial risk of investing in foreign currency bonds stems from uncertain *local* bond values due to interest rate risk as well as uncertain bond values in the chosen *base currency* due to exchange rate risk. As the market conditions change, portfolio rebalancing decisions are assumed to be made at successive time points. The related dynamic decision problem can be formulated as a multistage stochastic optimization problem. The optimal first period decision, which is of the ultimate interest, depends on the actions which would be taken in the future periods under realizations of uncertain events. Thus, the purpose of the model is to aid the decision maker to manage the portfolio ie. (re)position the available funds through periodical restructurings in a way that would yield the best return on the initial investment measured in the chosen base currency at a specified horizon date.

Driven by needs arising from the practising financial community, recent years have witnessed a rapid growth of the area of financial modeling where new computational techniques from management science are introduced to solve problems formulated from the framework of existing financial economic theories. Focusing specifically on bond portfolio problems, discrete scenario models for a single currency case have been proposed in eg. Bradley and Crane [1] who already in 1972 suggested dynamic linear programs for managing bond portfolios. Because of the numerical difficulties they faced in solving a linearly constrained nonlinear preference (utility) maximization problem, they maximized the expected terminal value of the portfolio and took risk aversion into account by posing additional constraints which limited the size of allowable capital loss per year. Solving the model was based on applying a decomposition algorithm on the equivalent deterministic problem with chosen outcomes for interest rates at each period. Consequently, this model was restricted to selecting only very few possible future course of events because of the fast growth of the problem size as the number of decision stages and number of allowable realizations are increased. More recently, Mulvey and Vladimirou [2] have applied a dynamic generalized network modeling approach to variety of asset allocation problems. Zenios [3] extended the Mulvey-Vladimirou model to include mortgage-backed securities. Hiller and Eckstein [4] proposed a stochastic risk/return efficient frontier model for selecting optimal static portfolios of fixed-income securities but in their model, rebalancing of the portfolio at successive points of time was not allowed.

The model to be presented in this paper uses a stochastic quasigradient method to solve a two-stage linearly constrained expected utility maximization problem. The second stage problem in the model is a nonlinear stochastic optimization problem. Thereby, the approach differs from the standard two-stage formulation. As this iterative algorithm uses an independent sample of scenarios at each iteration, a large number of scenarios can be considered on the way to the optimal solution.

2. Formulating the stochastic optimization model

2.1 Defining the instruments

We assume that there is a base currency in which the performance of the bond portfolio will be measured. The financial instruments to be included in the portfolio can be purchased from the international government bond markets. We will consider only straight bonds with a fixed coupon rate ie. bonds that provide the investor with a deterministic set of cash flows: prespecified coupon (interest) payments and the principal repayment upon maturity. Normally, the interest payment frequency is semi-annual in the United States and Japan, and annual in Europe. At any given time, the amount one has to pay for the bond is financially equal to the present value of the future cash flows related to the bond. Let $T=\{t_1,t_2,...,t_k\}$ be the points in time when payments of a given bond occur. Let $C(t_i)$ be the payment at t_i and let r_t define the term structure of interest rates. The theoretical price P of a bond at time $t=0$ expressed in local currency is

$$P=\sum_{t\in T}\frac{C(t)}{(1+r_t)^t} \tag{2.1}$$

where

$$C(t_i)=\begin{cases} C & \text{for } i<k \\ C+F & \text{for } i=k \end{cases} \tag{2.2}$$

Here C represents a fixed annual coupon interest payment and F represents the principal (face value) payment at the maturity date t_k. The interest rates r_t, used for discounting the cashflows, are stochastic. Therefore, the price of a bond is subject to interest rate risk. Traditionally, a bond's exposure to interest rate risk has been measured using the concept of duration, see eg. [5]. The effective duration measure can be calculated by shifting the yield curve by h to r_t+h, for all t, and differentiating the bond price with respect to h, at $h=0$.

We define the exchange rate $S_d(t)$ as the number of units in domestic currency for one unit of foreign currency at time t. Thus, the price of a foreign currency bond in the domestic currency becomes

$$P_d = PS_d(t) \tag{2.3}$$

Here $S_d(t)$ is stochastic which introduces the exchange rate risk to the problem. Thus, there are two sources of uncertainty which are interrelated. To illustrate this interaction, we can take an example from the foreign exchange markets in the fall of 1992 related to the Finnish, Swedish or Italian currencies. If the markets expect a country's currency to devaluate, the country's central bank will have to support its currency by raising interest rates. By doing this, the central bank hopes to attract market players to buy its currency instead of (short)selling it and thereby to restore the value of the currency.

2.2 Objective function and constraints

The use of the utility theory due to von Neuman and Morgenstern has become widely accepted in models of financial decision making under uncertainty. According to the utility theory, the choice between several alternatives with uncertain consequences is based on the expected utilities of these alternatives. Utility $u(w)$ in financial modeling is a monotonically increasing function of wealth w, as investors are assumed to prefer more wealth to less. The choice behavior of most investors is characterized by risk aversion which results in a strictly concave utility function. The question of expressing a person's attitudes towards risk as a mathematical function, is an area of research in itself explored by decision scientists and psychologists alike. Examples of attemps to empirically elicit the relationship between decision maker's risk preference and wealth can be found eg. in Gordon, Paradis, and Rorke [6] who employed portfolio games on decision makers to derive their utility functions.

A particular class of utility functions that is widely used in financial modeling is a family of HARA (hyperbolic absolute risk aversion) or equivalently, LRT (linear risk tolerance) utility functions [7]:

$$u(w) = \frac{1-\alpha}{\alpha}\left(\frac{\gamma w}{1-\alpha} + \beta\right)^{\alpha} \qquad , \beta \geq 0 \tag{2.4}$$

where $\gamma w/(1-\alpha) + \beta > 0$. The risk aversion can be expressed by Arrow-Pratt absolute risk aversion measure $A(w) = -u''(w)/u'(w)$. The absolute risk tolerance function $T(w)$ is obtained as an inverse of $A(w)$. For the utility function (2.4) we obtain

$$T(w) = \frac{w}{1-\alpha} + \frac{\beta}{\gamma} \tag{2.5}$$

which is linear with respect to level of wealth w. Risk aversion is decreasing for $\alpha < 1$ and increasing for $\alpha > 1$. From the general HARA function specification, several commonly used formulations are obtained with different parameter values: eg. quadratic with $\alpha = 2$, isoelastic with $\beta = 0$ and $\alpha < 1$, logarithmic with $\beta = \alpha = 0$, and negative exponential utility with $\beta = 1$ and $\alpha = -\infty$. Kallberg and Ziemba [8] have shown empirically that maximization of expected utility with different functional forms but similar Arrow-Pratt absolute risk aversion measures results in similar optimal portfolio allocations.

For our application, we select a negative exponential utility function which, in relation to (2.4), implies the following functional form:

$$u(w) = -e^{-\gamma w} \tag{2.6}$$

However, as we are most concerned about the relative change in the portfolio value ie. return on the invested capital from the horizon period, we incorporate the initial portfolio value to the utility function specification:

$$u(w) = -e^{-\gamma(\frac{w-w_0}{w_0})} \tag{2.7}$$

where w_0 and w represent the portfolio value at the outset and at the horizon date, respectively. This function specification has the property of constant absolute risk aversion where increasing the parameter γ implies more risk-averse choice behavior.

The optimization model will provide recommendations as to which bonds to hold in the portfolio given the existing and forecasted market conditions in the interest rate and foreign exchange markets. For a two-stage model, these recommendations are constrained by the current market value of the portfolio and by the next period market value of the portfolio determined by the interest rate and exchange rate realizations. Thus, the constraints to the decisions are simple budget constraints stating that the value of the portfolio immediately after rearranging cannot exceed the funds available for rearranging ie. current market value of the portfolio. As the allowable instruments are goverment bonds with good liquidity, the proxy for the market value of each bond at any given time is simply calculated as the present value of remaining cash flows converted to the base currency. At this stage, the transaction costs are omitted but theoretically, they could be taken into account in terms of bid and offer rates when pricing the bonds.

2.3 The problem

The horizon period is divided into two time periods which may be of unequal length. The allocation decision that is made now affects the terminal portfolio value through the recourse decision at the beginning of the second period. Let us introduce the following notation:

- i refers to a particular bond, $i = 1, ..., n$
- $x_0 = (x_0^i) \in R^n$ is the first stage decision vector: x_0^i is the number of bonds i to be included in the portfolio at time $t = 0$
- $x_1(\omega) = x_1^i(\omega) \in R^n$ is the second stage decision vector: $x_1^i(\omega)$ is the number of bonds i to be included in the portfolio at time $t = 1$ given first stage realization ω
- $P_0 = (P_0^i) \in R^n$ is the first stage price vector
- $P_1(\omega) = (P_1^i) \in R^n$ is the second stage price vector given first stage realization ω
- $P_2(\nu \mid \omega) = (P_2^i) \in R^n$ is the second stage ending price vector given first stage realization ω and second stage realization ν

- $C_j=(C_j^i) \in R^n$, for $j = 1, 2$, define cash flows of bonds i in period j prolonged to the end of period j employing interest rate from the beginning of period j. Note that $C_2^i=C_2^i(\omega)$ is defined conditional on ω.
- b is the initial wealth

The objective of the model is to maximize the expected utility of the portfolio value at the horizon date:

$$\max_{x_0} f(x_0) = E_\omega v(x_0, \omega) \tag{2.8}$$

subject to

$$\begin{aligned} P_0 x_0 &\leq b \\ x_0 &\geq 0 \end{aligned} \tag{2.9}$$

where E_ω refers to expectation over first stage realization ω and

$$v(x_0, \omega) = \max_{x_1(\omega)} E_\nu u\Big((P_2(\nu \mid \omega) + C_2(\omega))x_1(\omega)\Big) \tag{2.10}$$

subject to

$$\begin{aligned} P_1(\omega)x_1(\omega) &\leq (P_1(\omega) + C_1)x_0 \\ x_1 &\geq 0 \end{aligned} \tag{2.11}$$

where E_ν refers to expectation over second stage realization ν.

Let us note that our second stage decisions are solutions of stochastic optimization problems. In conventional formulations, after observation of ω, all relevant information is revealed and at the second stage we are confronted with a deterministic problem. The second stage problem is also nonlinear in contrast to the usually discussed case.

3. Interest rate and exchange rate dynamics

Given that both the interest rate movements and exchange rate movements in the market are stochastic, highly unpredictable, and heavily affected by market emotionalism and expectations, this uncertainty must be explicitly modeled. As the bond values are functions of interest rates and exchange rates, formulation of the portfolio optimization problem requires assumptions about the dynamics of these random variables. There are several alternative modeling approaches for interest rate and exchange rate dynamics to be found in theoretical literature of monetary economics. Thus, the choice of the most appropriate one is by no means trivial but poses an important modeling question.

3.1 Continuos time approaches

Continuous time models are widely used in many fields of economics and especially in finance. A basic form for a stochastic variable z that follows a generalized Wiener process with drift is

$$dz = \mu dt + \sigma dW(t) \tag{3.1}$$

where $W(t)$ is a standard Wiener process and μ and σ are the drift and volatility parameters (here constants) respectively. Classical applications to model the dynamics of the term structure of interest rates include the Vasicek and Cox-Ingersoll-Ross, and Brennan-Schwartz models (for a summary, see eg. [10]). They differ in their attempts to specify the "correct" volatility function σ. Perhaps the most widely known applications using stochastic differential equations for modeling exchange rate dynamics are the pricing formulas for foreign exchange options which are based on the pioneering work of Black and Scholes on stock option pricing.

In the recent academic literature on pricing of foreign currency options under stochastic interest rates Amin and Jarrow [9] start off by assuming that both interest rate and exchange rate movements over time can be modeled as diffusion processes. Given that interest rates of different economies are correlated as are also these interest rates and the respective exchange rates, these relationships need to be made explicit. The correlations between the variables can be modeled by employing independent random factors, here standard Wiener processes, which the correlated variables have in common. In the following, the formulation of spot interest rate at time t for maturity at time T consists of a non-stochastic trend component μ_d and two stochastic variance components σ_{di}, i=1,2. The variance components are assumed to depend on the time of observation t, the "term" of the interest rate T, and the current level of the interest rate $r_d(t,T)$. Furher, it is assumed that there are two standard Wiener processes $W_i(t)$ associated with the variance componens that influence the behavior of the entire term structure. Often these factors are interpreted as short-term and long-term or inflation factor. To summarize, it is assumed that for each currency, the domestic term structure of spot interest rates curve $r_d(t,T)$ follows a stochastic differential equation of the form

$$dr_d(t,T) = \mu_d(t,T)dt + \sum_{i=1}^{2} \sigma_{di}(t,T,r_d(t,T))dW_i(t) \tag{3.2}$$

A long-term factor, $W_2(t)$, is used to model the prevailing correlation between interest rate movements in the two countries and will therefore also appear in the formulation of the foreign interest rate term structure. The foreign interest rate term structure $r_f(t,T)$ is thus assumed to follow a process determined by

$$dr_f(t,T) = \mu_f(t,T)dt + \sum_{i=2}^{3} \sigma_{fi}(t,T,r_f(t,T))dW_i(t), \tag{3.3}$$

with $\sigma_{f2}W_2$ being the long-term and $\sigma_{f3}W_3$ being the short-term random factor for foreign interest rates.

As before, define the exchange rate $S_d(t)$ as the number of units in domestic (base) currency for one unit of foreign currency at time t. Then the diffusion process for the spot foreign exchange rate is

$$dS_d(t) = \gamma_d(t)S_d(t)dt + \sum_{i=1}^{4} \delta_{di}(t)S_d(t)dW_i(t) \tag{3.4}$$

Here, the trend and volatility components are γ_d and δ_{di}, respectively. Note that all the three previously introduced random factors are present to account for correlation between domestic and foreign interest rates and the respective exchange rate. Additionally, an independent self-induced random disturbance, $W_4(t)$, is allowed to affect the exchange rate dynamics. Given these processes for interest rates and exchange rates, stochastic calculus can be used to derive the dynamic process for the portfolio value, as indicated in [9].

Despite its attractive generality, the use of this kind of thoretical model for the behavior of underlying stoachastic variables poses some difficulties for practical application. Most importantly, the estimation of the drift and volatility parameter function is made difficult by the simple notion that stochastic differential equations study *continuous* time while the data is observed at *discrete* time intervals. We are faced with a so-called aliasing problem [10], which states that if the process is observed only at discrete intervals, it is impossible to identify the stochastic process that has generated the data set, especially if there are unobservable state variables (e.g. conditional variance) in the process. However, this modeling approach reveals the complexity of the factors of uncertainty pertaining to the optimal fund allocation problem. Furthermore, from the diffusion model approach we adopt the formulation of interdependence of the underlying stochastic variables when developing a working discrete time model for forecasting purposes.

3.2 Discrete time approaches

From an empiricist's perspective, instead of stochastic differential equations, it is more tempting to use discrete time stochastic difference equations to model the stochastic behavior of changes in interest rates and exchange rates. Many econometric studies conducted on the stochastic behavior of financial price changes have rejected the assumptions of normality and constant conditional variance of variable changes and concluded that in high-frequency financial time series, the volatility seems to be clustered and forecastable. One of the most popular approaches to model the heteroskedasticity in time series data has been the use of ARCH (Autoregressive Conditionally Heteroscedastic) models [11]. This class of dynamic models of time dependant conditional variance was introduced by Engle in 1982. In the original ARCH(q) specification, the conditional variance h_t of the error term is calculated as a weighted average of q past squared forecast errors ϵ_t:

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 \tag{3.5}$$

where the α_i's are parameters to be estimated and q is the order of the ARCH process.

In the literature, various extensions to the original ARCH specification have been proposed. One of the most widely used is the GARCH(p,q) (Generalized ARCH) model which specifies the conditional variance as follows:

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} \tag{3.6}$$

To estimate the parameters α_i and β_i in the above equation a conventional maximum likelihood method can be applied. The resulting distribution for the forecast error ϵ_t is Gaussian $\mathcal{N}(0,h_t)$.

An extensive survey on ARCH modeling in finance by Bollerslev, Chou, and Kroner [12] lists several studies that provide evidence on very significant ARCH effects on interest rates. Also, findings are reported from several studies according to which main currencies behave as martingale processes with leptocurtic distributions and conditional heteroskedasticity.

As was discussed in the previous section, the interest rate and exchange rate variables cannot be treated as independent stochastic variables but, their covariances with each other must be taken into account. Therefore, it becomes necessary to analyze the stochastic process of the system of variables in a multivariate context. The formulation of a multivariate GARCH model is equivalent to the one given in equation (3.6) but now the model pertains to the conditional variance-covariance matrix. Although theoretically appealing, this approach suffers from the large number of parameters that need to be estimated. One simplification is to assume that the conditional covariance matrix varies over time but that the conditional correlations between forecast errors remain constant. Empirical studies have found this to be a reasonable working hypothesis.

Literature concerning the bridging of the continuous time stochastic differential equation systems and discrete time stochastic difference equation systems is fairly limited. In his paper, Nelson [13] shows that if the true model is a diffusion model without jumps, then the discrete time variances are consistently estimated by a weighted average of past forecast errors as in GARCH(1,1) formulation.

Other stochastic model specifications that could be used for interest rate formulation include the discrete binomial lattice approach (see eg.[5]) which can be used to generate arbitrage-free interest rate scenarios that are consistent with the currently observable interest rate term structure. This type of modeling approach is adopted in portfolio optimization applications reported eg. in Zenios [3] and Hiller and Eckstein [4]. However, there are also empirical studies arguing that especially long-term interest rates are too volatile to be forecasted based on the expectation theory with a constant liquidity premium [12]. Therefore, no unambiguous answer can be given as to how to correctly model the dynamics of the underlying stochastic variables.

3.3 Simplifying assumptions on stochasticity

Defining the most appropriate way to model the stochastic behavior of interest rates and exchange rates for bond portfolio optimization will be a subject of further research, and for now, we will settle on simple Wiener process specifications with time-dependent variance. (In the following, stochastic variables refer to logarithmic changes in interest rates and exchange rates.) We assume that each stochastic variable z follows a process determined by

$$dz = \sigma dW(t) \tag{3.7}$$

where $W(t)$ is a standard Wiener process and σ is the volatility parameter. Thus, the change in each variable is normally distributed $\mathcal{N}(0,\sigma dt)$. We use a sample variance to estimate parameter σ from historical data. Note that the variance depends on the decision stage through dt. We also make the simplifying assumption that the correlations between the changes in stochastic variables remain constant over time but that the variances depend on time in the manner described by (3.7) above. Therefore, the distribution of the changes in stochastic variables is assumed to be multivariate normal $\mathcal{N}(0,\Sigma)$ where Σ is the time-dependent variance-covariance matrix. Such distributions are employed to generate first and second stage realizations for interest rates and exchange rates. These realizations are then used to calculate the market prices for the bonds, according to (2.1)-(2.3).

Let us elaborate further on generating the interest rate and exchange rate outcomes which are consistent with the assumed distribution. Let P denote the correlation matrix estimated from historical data for changes in stochastic variables and let D be the diagonal matrix for standard deviations of these variables at a particular stage, say at stage one. Then, the variance-covariance matrix Σ is given by

$$\Sigma = DPD. \tag{3.8}$$

The correlation matrix P is positive definite and symmetric so that Cholesky factorization can be applied to obtain a lower triangular matrix L such that $LL^T{=}P$ (see eg. [16]). We generate a vector y, having equally many components as there are stochastic variables. Each component of y is drawn independently from a univariate normal distribution $\mathcal{N}(0,1)$. Now, the vector of stochastic variable outcomes for the first stage realization is obtained by adding DLy to the vector of current levels of interest rates and exchange rates. This procedure is consistent with maintaining the assumed variance-covariance structure as

$$E(DLy(DLy)^T = E(DLyy^TL^TD) = DLL^TD = DPD = \Sigma. \tag{3.9}$$

The second stage realizations are generated conditional to the first stage realization according to a martingale process. The first stage realization determines the base values for the stochastic variables to which the *changes* in these variables are applied, as generated from the relevant multivariate normal distribution. The random vector generation for the second stage is done as above with Σ now referring to the second stage covariance matrix.

4. Stochastic optimization

Stochastic optimization models with recourse were first formulated in mid-50's by Dantzig and Beale for linear programs with random coefficients. Ever since the introduction of stochastic dynamic programming models, there has been an ongoing search for efficient solution algorithms. One approach has been to solve the equivalent deterministic problem by taking advantage of the structure of the problem. However, deterministic approaches are faced often with the problem of extremely large problem sizes, as the number of scenarios and number of dynamic stages increase. Therefore, it becomes necessary to carefully select the critical scenarios to be considered by the model. As of today, there is no generally accepted criteria as to how this should be done. Another approach is to use sample based methods which take advantage of the structure of a dynamic stochastic problem but, which also employ Monte Carlo sampling techniques to master the inherent stochasticity. One example of this type of methods is the dual (Benders) decomposition method combined with importance sampling that Dantzig and Glynn [14] proposed for solving stochastic linear programs. In the stochastic quasigradient method by Ermoliev

[15] adopted for our application, the sub-problems are approximated by taking independent samples of scenarios. This yields approximations of subgradients to be used for updating the first stage variables.

4.1 The stochastic quasigradient method

The problem at hand involves stochasticity in form of uncertain bond values as an outcome of different realizations for interest rates and exchange rates. As formulated in (2.8)-(2.11), both the coefficients of the objective function and the coefficients in the second stage budget constraints are affected by these uncertainties. However, the optimal allocation of funds must be found now, before observing any realization for interest rates or exchange rates.

As a solution strategy, a stochastic quasigradient method developed by Ermoliev [15] will be used. This method does not assume differentiability of the objective function, existence of gradients, or ability to evaluate the function value or its gradient exactly. The method uses statistical estimates for function values and subgradients. Our objective function is a concave function of decision variables. The calculation of gradients for each possible realization is not possible due to the infinite number of events in the sample space and therefore, some kind of approximation method is needed. The stochastic quasigradient method applied to the problem (2.8)-(2.11), generates at each iteration a single scenario ω of the first stage realization at random. The second stage problem (2.10)-(2.11) is approximated by a sample mean, conditional on ω, by taking a random sample N of second stage realizations $\{\nu \mid \omega\}$. We solve the resulting deterministic sub-problem and obtain $x_1(\omega)$. On the basis of this solution, we derive an estimate of the subgradient of the function f (stochastic quasigradient) and change the current approximate solution of the first stage in this direction. We define the direction vector $\xi(x_0, N \mid \omega)$ as follows:

$$\xi(x_0, N \mid \omega) = \lambda(x_0, N \mid \omega) P_1(\omega) \tag{4.1}$$

where $\lambda(x_0, N \mid \omega)$ is the optimal dual multiplier of the budget constraint in the second stage optimization restricted to sample N, for a given x_0. Now, the fundamental notion within the method is the fact that for the direction vector $\xi(x_0, N \mid \omega)$, as $\mid N \mid \to \infty$, the following conditional expectation statement holds:

$$E_\omega\{\xi(x_0, N \mid \omega)\} = f^N_{x_0}(x_0), \quad f^N(x_0) \to f(x_0). \tag{4.2}$$

where f_{x_0} denotes a subgradient of the function $f(x_0)$ and E_ω refers to expectation over ω. In other words, *on the average*, the direction vector $\xi(x_0, N \mid \omega)$ belongs to the subgradient set of the function $f^N(x_0)$ converging to $f(x_0)$. Such type of situation is treated within a non-stationary optimization framework. Since f_{x_0} is a subgradient and since ξ is, due to the approximate solution of the second stage problem, a *biased* estimate of f_{x_0}, there is no guarantee to decrease the expectations. Noting this, a monotonic increase in the objective function cannot be guaranteed as we update x_0 in the direction $\xi(x_0, N \mid \omega)$ at each successive iteration but, the convergence takes place with probability 1. We will carry out experiments to find a practical sample size $\mid N \mid$ for our problem. This is easy to obtain simultaneously with iterations when a certain stability of the current approximate solution is reached.

To formulate the procedure in terms of an algorithm, we start iteration $s = 0$ with the original portfolio allocation. In each iteration s, $s = 0, 1, ..$, we observe the first stage realization ω, solve for the approximated second stage optimization problem in (2.10)-(2.11) by employing a random sample N of second stage realizations ν, *conditional* on realization ω. The dual solution $\lambda(x_0, N \mid \omega)$ is used to update x_0^s as follows:

$$x_0^{s+1} = [x_0^s + \rho^s \xi(x_0, N \mid \omega)]_X \tag{4.3}$$

where $X = \{x_0 \mid P_0 x_0 \leq b, \ \ x_0 \geq 0\}$ is employed in orthogonal projection $[\,\cdot\,]_X$, and ρ^s is the step size satisfying

$$\rho^s \to 0, \ \ \sum_{s=0}^{\infty} \rho^s = \infty, \ \ \text{and} \ \ \sum_{s=0}^{\infty} (\rho^s)^2 < \infty. \tag{4.4}$$

Next, we observe a new ω, approximate the arising second stage subproblem by a finite number of observations and so on. For proof of convergence of the algorithm, see Ermoliev [15], including the case of non-stationary objective functions. In our case, we restrict the discussion to only a finite number of second stage scenarios, which also fits the existing proof. Otherwise, we would need an appropriate statement in terms of ϵ-subgradients or non-stationary optimization [15].

Stochastic quasigradient methods tend to reach the neighborhood of optimal solutions fairly fast but the tail for obtaining a highly accurate solution can be long. Considering the nature of the bond portfolio allocation problem, which aims at producing recommendations to support practical decision making, getting reasonably close to optimal solutions will be sufficient.

4.2 Confidence intervals for the optimal value

An important question concerns the selection of a suitable stopping criterion for the stochastic quasigradient method. In order to make judgements on the quality of a near-optimal solution we will develop probabilistic bounds for the optimal objective function value. Let x_0^* denote an approximative optimum. Based on concavity of our objective function, the following inequality holds:

$$f(x_0) \leq f(x_0^*) + \nabla f(x_0^*)(x_0 - x_0^*) \ \ \forall x_0. \tag{4.5}$$

Let us denote $\Delta = x - x_0^*$. We can now approximate the right-hand side of the above inequality by performing a large number of evaluation iterations at x_0^* to obtain a sample mean estimate a for the objective function value $f(x_0^*)$ and a sample mean estimate b for the gradient $\nabla f(x_0^*)$. We use crude Monte Carlo sampling for evaluation. (More sophisticated sampling methods will be considered at a later stage of our research.) Define the random vector $r = (a, b)$. The variance-covariance matrix estimate R, for r, is obtained as sample variances and covariances. The variance of the estimate for the *upper bound* in (4.5) is then given by

$$\sigma_f^2(\Delta) = (1, \Delta) R \begin{pmatrix} 1 \\ \Delta \end{pmatrix} \tag{4.6}$$

As we employ a large sample of independent evaluations, the estimated bound $a + b\Delta$ on the objective function is approximately normally distributed.

The first stage solution vector x_0^* is feasible, so that $f(x_0^*)$ provides a *lower bound* for the true optimum value. However, because of the stochasticity involved in estimating $f(x_0^*)$, we can only express the lower bound in probabilistic terms. Thus, our lower bound for the true optimal value at eg. 97.5 % confidence level becomes

$$\underline{f} = a - 1.96\sigma_f(0) \tag{4.7}$$

Similarly we obtain an *upper bound* for the optimal value at a 97.5 % confidence level as the optimal value of the following problem:

$$\max_{\Delta} f(\Delta) = a + b\Delta + 1.96\sigma_f(\Delta) \tag{4.8}$$

subject to (4.6) and

$$\begin{aligned} P_0\Delta &= 0 \\ x_0 + \Delta &\geq 0 \end{aligned} \tag{4.9}$$

As the objective function in (4.8) is a convex function defined on the compact convex set, we find the maximum for (4.8)-(4.9) in vertices of the budget simplex [17].

4.3 Algorithm

All the above considerations can be now summarized in the following outline of the stochastic quasigradient algorithm for the portfolio allocation problem:

- Step 0. To initiate, set s=0 and x_0^s equal to the initial portfolio allocation
- Step 1. Generate one scenario ω for the first stage and use this scenario to calculate the bond prices P_1 in the base currency
- Step 2. Generate a sample N of second stage realizations conditional on ω and solve the approximative second stage problem with budget $P_1 x_0^s$ using an interior point method.
- Step 3. Update the first stage decision variables using the price vector P_1 and the optimal dual multiplier λ of the second stage budget constraint:

$$x_0^{s+1} = [x_0^s + \rho^s \lambda P_1]_X \tag{4.10}$$

  where $\rho^s = \rho^0/s$ and $[\cdot]_X$ refers to projection on the set $X = \{x_0 \mid P_0 x_0 \leq b,\ \ x_0 \geq 0\}$
- Step 4. Perform a probabilistic optimality test (of Section 4.2) at every κth iteration. If the test is passed, stop; otherwise, return to Step 1.

5. Computational experience

For setting up a test application, the set of financial instruments was selected out of the most liquid government bonds denominated in three currencies: U.S. dollar, Deutschmark, and Japanese yen. For each currency, bonds were selected to represent approximately maturities of three, five, and ten years, as it was seen desirable to choose bonds with different interest rate sensitivities. Thus, the following computational results are based on the test application which determines the optimal funds allocation between nine government bonds. On each bond, the following information was input as parameters for the pricing module: currency of denomination, coupon rate and payment schedule, face value, and maturity date. The initial prices were calculated by discounting the cash flows by prevailing three month market interest rate by assuming a flat yield curve for simplicity.

The chosen horizon period of six months was divided into two periods the first one being one month and the second one five months. The notion that it is increasingly difficult to make predictions on interest rate and exchange rate behavior further out to the future, gets taken into account as a greater variance for the variable changes is employed for the second period.

Several simulations were run with the model to test how changes in different parameter values affect the performance of the algorithm. It was found particularly interesting to study how the number of scenarios generated at each iteration for the second stage optimization problem influenced the ability of the method to find a near-optimal solution. Secondly, we wanted to solve the model using different values for the risk aversion parameter to see the impact of stochasticity when the curvature of the objective function varies. Finally, we were also curious to find out about the role of the choice for the initial step size. In all cases, we tested the dependence of precision (confidence interval) as a function of iterations. All the simulations were run on a HP9000/720 computer.

In the following tables, the "Iterations" column refers to the number of main loop iterations which update the first stage decision variables. The "Objective function" column refers to the estimated objective function value obtained by evaluating the objective function after the given number of iterations. Columns "Upper error" and "Lower error" refer to the relative error of the objective function at 95% confidence level expressed in percent terms (relative to the absolute value of the objective function estimate). The "Return/risk" column contains the expected annualized return and the annualized standard deviation of the return for the optimal portfolio. The calculation of the mean objective function value, the relative errors as well as the calculation of expected return and its standard deviation were based on 10000 evaluation iterations each using 100 scenarios for the second stage. The "Distance" column indicates the relative distance between the obtained solution, say x_0, and the solution $\hat{x}_0$ obtained after 10000 main loop iterations. Relative distance of x_0 and $\hat{x}_0$ is defined as $\| x'_0 - \hat{x}'_0 \| \,/\, \| \hat{x}'_0 \|$, where x'_0 and $\hat{x}'_0$ are vectors of value distributions; i.e. x_0 and $\hat{x}_0$ with components scaled by bond prices.

Finally, the column "CPU time" refers to the CPU time used in seconds for the optimization iterations. (This is reported exactly for each simulation using 1000 main loop iterations and for others, the figure is approximate based on the observation that the used CPU time is proportional to the number of main loop iterations.)

Let us first look at the "base case" where the risk aversion parameter γ in (2.7) is 20 and the number of scenarios used in the second stage optimization is 100.

Iterations	Objective function	Upper error (%)	Lower error (%)	Return/ risk	Distance (%)	CPU time (s.)
100	-0.6527	1.2	0.3	4.7/2.5	24.0	6
200	-0.6522	2.2	0.3	4.6/2.2	15.7	12
500	-0.6520	3.1	0.2	4.5/2.0	9.1	28
1000	-0.6515	2.3	0.2	4.6/2.1	8.1	56
2000	-0.6513	2.3	0.2	4.6/2.1	5.8	112
5000	-0.6509	2.9	0.2	4.6/2.0	2.1	280
10000	-0.6498	3.0	0.2	4.6/2.0	0	560

As can be noted, the optimal value of the objective function stabilizes fairly quickly reaching the vicinity of the optimum after about 1000-2000 main loop iterations. The relative error bounds for the objective function value are generally fairly small even though the upper bound is calculated in a very conservative way based on gradient approximation. The expected annualized return stabilizes at 4.6 % and annualized standard deviation of the return at 2.0 %. For the otherwise equivalent case but with the number of second stage scenarios being only 10, the following results were obtained:

Iterations	Objective function	Upper error (%)	Lower error (%)	Return/ risk	Distance (%)	CPU time (s.)
100	-0.6550	4.6	0.3	4.4/2.3	17.3	1
200	-0.6547	2.8	0.3	4.5/2.6	16.3	2
500	-0.6547	2.6	0.3	4.5/2.4	13.1	4
1000	-0.6536	2.6	0.3	4.5/2.4	9.8	8
2000	-0.6530	1.5	0.3	4.5/2.3	6.2	16
5000	-0.6527	1.7	0.3	4.6/2.2	3.1	40
10000	-0.6510	1.2	0.3	4.6/2.2	0	80

The convergence to the optimal solution is now slower than in the "base case" which is why we prefer to use at least 100 second stage scenarios at each iterate. Another reason for using a larger sample size is to decrease the impact of statistical biases in the gradient estimates. To evaluate whether 100 scenarios is a sufficiently large sample size we performed the equivalent optimization by employing 1000 scenarios in the second stage optimization:

Iterations	Objective function	Upper error (%)	Lower error (%)	Return/ risk	Distance (%)	CPU time (s.)
100	-0.6557	1.9	0.3	4.7/3.0	25.6	53
200	-0.6547	1.8	0.3	4.7/3.0	23.3	106
500	-0.6547	2.6	0.3	4.5/2.5	12.6	267
1000	-0.6513	2.7	0.3	4.6/2.4	8.9	534
2000	-0.6526	1.7	0.3	4.6/2.4	8.2	1068
5000	-0.6502	1.6	0.3	4.7/2.2	3.4	2670
10000	-0.6507	2.6	0.2	4.6/2.2	0	5340

As the large sample size does not lead to any visible improvement in the performance, we will keep the sample size at 100 for the remaining simulations.

Secondly, the effect of the objective function concavity was analyzed. We first studied a less risk-averse case by decreasing the value of parameter γ from 20 to 10; the number of second stage scenarios as well as other parameters were kept as defined in the "base case" run. This provided us with the following results:

Iterations	Objective function	Upper error (%)	Lower error (%)	Return/ risk	Distance (%)	CPU time (s.)
100	-0.7918	1.8	0.1	4.6/2.2	6.6	5
200	-0.7919	2.0	0.1	4.5/2.1	4.2	10
500	-0.7920	2.1	0.1	4.5/2.0	3.8	27
1000	-0.7916	1.9	0.1	4.6/2.1	2.4	54
2000	-0.7914	1.9	0.1	4.6/2.1	1.8	108
5000	-0.7912	2.1	0.1	4.6/2.1	0.8	270
10000	-0.7904	2.2	0.1	4.6/2.2	0	540

Next, we studied the problem from a considerably more risk-averse point of view. The value of the risk aversion parameter γ was increased to 50 which produced the following results:

Iterations	Objective function	Upper error (%)	Lower error (%)	Return/ risk	Distance (%)	CPU time (s.)
100	-0.3867	3.0	0.6	4.5/2.0	23.8	6
200	-0.3832	5.4	0.5	4.4/1.5	7.9	12
500	-0.3821	11.2	0.4	4.3/1.3	3.7	30
1000	-0.3814	3.2	0.4	4.4/1.5	2.2	61
2000	-0.3812	4.4	0.4	4.4/1.4	0.6	122
5000	-0.3809	7.2	0.4	4.4/1.4	0.9	305
10000	-0.3799	6.6	0.4	4.4/1.4	0	610

Based on the above two runs it can be concluded that the less well-behaving the objective function becomes, the more the impact of stochasticity gets emphasized, as expected. With increasing concavity, the relative error bounds for the objective function value also increase which can be explained by a greater variance in the expected values for the objective function value and the gradient estimate. For the more risk-averse case, the optimal portfolios have now notably smaller standard deviation but also lower return which is consistent with the concept of risk-return trade-off.

Finally, the role played by the choice for the initial step size was analyzed. All the above simulations were run with the initial step size ρ^0 being 10^3. We solved the model by employing again 100-10000 main loop iterations with varying initial step sizes. The simulations lead us to conclude that the algorithm is fairly insensitive to the selection of initial step size as long as it is in the range of 10^2-10^4.

There is at least one more interesting observation to report: it appeared that there was some amount of variation in the optimal first stage variable values between simulation runs even when the estimates for the objective functions were almost the same. This lead us to conclude that the set of near-optimal solutions is relatively large. This notion was verified by taking convex combinations of different "equally optimal" allocations and evaluating the resulting portfolio for its expected utility. Indeed, the expected optimum value for the convex combination was at least no worse than the expectation for any of the "equally optimal" portfolios alone.

6. Conclusions and subjects for future research

We have discussed the formulation and solution of the decision problem related to the optimal management of a multi-currency bond portfolio. We formulated the problem as a two-stage stochastic optimization problem where the stochasticity stems from interest rate and exchange rate uncertainty. We implemented the model and used a stochastic quasigradient method as the solution method. For optimality tests, probabilistic error bounds were developed.

It turned out that the stochastic quasigradient method was efficient in quickly providing a solution that lies in the vicinity of the true optimal solution. Even a small random sample of second stage scenarios at each iterate was enough to bring about a quick convergence to close to the optimum values. However, each simulation run indicated that improved accuracy could only be achieved at an increasingly high cost measured in terms of solution time. Roughly speaking, it was felt that satisfactory precision for this problem was obtained by using about 1000-2000 independent main loop iterations each using 100 second stage scenarios generated conditionally on the first stage scenario. This took less than two minutes to solve. To re-emphasize, the accuracy was felt to be sufficient bearing in mind the practical decision supporting function of our application.

It would be interesting to investigate, whether another kind of sampling method, eg. importance sampling, would result in improved performance of the stochastic quasigradient method. We also plan to implement the model by using the saddle point algorithm developed for solving the deterministic equivalent problem, and compare efficiency with the current approach. Another area of further research is the integration of the tools from econometrics and finance to introduce an improved formulation for stochasticity underlying the model.

References

[1] S.P. Bradley and D.B. Crane, "A Dynamic Model for Bond Portfolio Management", *Management Science* Vol. 19, No. 2 (1972).

[2] J.M. Mulvey and H. Vladimirou, "Stochastic Network Programming for Financial Planning Problems", *Management Science* Vol. 38, No. 11 (1992).

[3] S.A. Zenios, "Massively Parallel Computations for Financial Planning under Uncertainty", in J.P Mesirov (Ed.), *Very Large Scale Computation in the 21st Century* SIAM, Philadelphia, PA, 1991.

[4] R.S. Hiller and J. Eckstein, "Stochastic Dedication: Designing Fixed Income Portfolios Using Massively Parallel Benders Decomposition", *Management Science* Vol. 39, No. 11 (1993).

[5] T.S.Y. Ho, *Strategic Fixed Income Investment*, Dow Jones-Irwin, 1990

[6] M.J. Gordon, G.E. Paradis, and C.H. Rorke, "Experimental Evidence on Alternative Portfolio Decision Rules", *American Economic Review* Vol. 62 (1972).

[7] J. Ingersoll, *Theory of Financial Decision Making*, Rowman & Littlefield, 1987.

[8] J.G. Kallberg and W.T. Ziemba, "Comparison of Alternative Utility Functions in Portfolio Selection Problems", *Management Science* Vol. 29, No. 11 (1983).

[9] K.I. Amin and R.A. Jarrow, "Pricing Foreign Currency Options Under Stochastic Interest Rates", *Journal of International Money and Finance* No. 10 (1991).

[10] Y. Ait-Sahalia, "Non-parametric Pricing of Interest Rate Derivative Securities", Working Paper, University of Chicago, Graduate School of Business, 1993.

[11] R.F. Engle, "Statistical Models for Financial Volatility", *Financial Analysts Journal* Jan.-Feb. 1993.

[12] T. Bollerslev, R.Y Chou, and K.F. Kroner, "ARCH Modeling in Finance", *Journal of Econometrics* No 52 (1992).

[13] D.B. Nelson, "ARCH Models as Diffusion Approximations", *Journal of Econometrics* No 45 (1990).

[14] G.B. Dantzig and P.W. Glynn, Parallel Processors for Planning Under Uncertainty", *Annals of Operations research* No 22, 1990.

[15] Yuri Ermoliev, "Stochastic Quasigradient Methods", *Numerical Techniques for Stochastic Optimization*, Yu. Ermoliev and R.J-B Wets (Eds.), Springer-Verlag, 1988.

[16] R.Y. Rubinstein, *Simulation and the Monte Carlo Method*, John Wiley & Sons, 1981.

[17] D.G. Luenberger, *Introduction to Linear and Nonlinear Programming*, Addison-Wesley Publishing Company, 1973.