Nonlinear Optimization Methods with Financial Applications

John R. Birge and Zhen Liu, for Ziena (January 2007)

Financial applications have a long history of including optimization, starting with Markowitz's origin of the quadratic optimization model for determining an efficient portfolio to minimize variance for a given return. Portfolio optimization continues to be an active area with most applications focused on linear and quadratic optimization. General nonlinear optimization arises in this area, however, as models begin to acknowledge and capture the nonlinearity, asymmetry, and non-normality associated with returns in practice. In addition, complex financial products often involve a variety of nonlinear relationships that lead to nonlinear optimization in parameter estimation, tracking, and hedging. Credit instruments and their risk management also introduce nonlinearities that are difficult to include in linear or quadratic models. This note summarizes some of these financial applications and the role that nonlinear optimization methods can play in their solution.

1 Parameter Calibration

1.1 Volatility Estimation

Volatility describes how much the stock prices, interest rates, etc., change up and down around their means. The discussion in this subsection is largely based upon the work of Coleman, Kim, Li and Verma in [1] and [2].

In the famous Black-Scholes-Merton (BSM) equation for pricing European options, volatility is assumed to be constant. For a call option with a strike price of K and current share price S, the volatility implied by the BSM formula for an *at-the-money* ($S = K$) option is typically lower than the volatility for *in-the-money* ($S > K$) or *out-of-the-money* ($S < K$), options. This empirical evidence then exhibits a *smile structure* that suggests that volatility changes over share price and time. Finding mis-priced options or pricing an option that is not in the market generally requires knowledge of how volatility is a function of these parameters. (While more sophisticated stochastic volatility models attempt to explain the volatility smile in other ways, in practice, traders tend to use the basic BSM model with adjustments for volatility depending on price and time.)

In order to specify a volatility function $\sigma(S,t)$ with respect to stock price and time, we assume that we are given market prices of n call options with strikes K_j and maturities T_j. Using the closing prices C_j for the options under consideration, we can solve the following optimization problem

$$\min_{\sigma(S,t)} \Sigma_{j=1}^{n} (C(\sigma(S,t), K_j, T_j) - C_j)^2 \tag{1}$$

where $C(\sigma(S,t), K_j, T_j)$ is the call option price with respect to $\sigma(S,t), K_j$ and T_j. The price of a European call option is given by the Black-Scholes-Merton formula,

$$C = S\Phi(d_1) - Ke^{-rT}\Phi(d_2), \tag{2}$$

where

$$d_1 = \frac{log(\frac{S}{K}) + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \tag{3}$$

$$d_2 = d_1 - \sigma\sqrt{T} \tag{4}$$

and the option expires at time T with strike K. $\Phi(.)$ is the cumulative distribution function for the standard normal distribution, r is risk-free rate and σ is the volatility of the underlying security.

This is a nonlinear least square problem since the function $C(\sigma(S,t), K_j, T_j)$ depends nonlinearly on the decision variable $\sigma(S,t)$. Volatility is perhaps the most common of such parameters to estimate in this way. Other parameters (such as mortgage prepayment rates, credit spreads for different classes of debt, and higher price moments) lead to similar nonlinear least square formulations.

2 Asset-Liability Management (ALM) with Convex Transaction Costs or Nonlinear Utilities

In [4], Birge develops a multistage nonlinear stochastic program for the ALM problem as follows. This problem pertains to any institution (or individual) that invests in a set of assets to meet a stream of liabilities over time. Large users of this technology include pension funds and insurance companies. Recently, models have also been developed for individuals, particularly for retirement planning.

In general, a set of scenarios s is given (or simulated) across time.

The decision variables can then be given as:
$x_t(k,s)$: position of asset k at time t in scenario s,
$u_t^+(k,s)$: value of asset k bought at time t in scenario s,
$u_t^-(k,s)$: value of asset k sold at time t in scenario s.

The parameters then include the scenarios s, assets k, liabilities L, transaction costs (α^+, α^-), discount factor β, probabilities p, and utility function c.

The objective is to maximize the cumulative expected discount utility

$$max \sum_t \sum_s p_{st}\beta_{st}c_t(u_t(s), x_t(s))$$

such that for all s, k, t

$$r_t(k,s)x_t(k,s) + (1-\alpha^+)u_t^+(k,s) + (1+\alpha^-)u_t^-(k,s) = x_{t+1}(k,s) \tag{5}$$

$$\sum_k(-u_t^+(k,s) + u_t^-(k,s)) = L_t(s), u_t^+ \geq 0, u_t^- \geq 0 \tag{6}$$

$$u_t^{+/-}(k,s) = u_t^{+/-}(k,s^{'}) \text{ if } s, s^{'} \text{ have same history at } t \tag{7}$$

When the utility function c is nonlinear such as a power or exponential utility, or the transaction costs are convex with respect to value of assets $u_t^+(k,s), u_t^-(k,s)$, we have a nonlinear stochastic program. While the objective functions can be linearized, the solutions that correspond to any remaining linear pieces (i.e., very high or low values) become unrealistic since they do not account for risk. A consistent nonlinear objective can avoid these problems. These problems can also include additional risk constraints (such as those involving probability of achieving a goal or avoiding a trap) that are typically nonlinear (and often nonconvex) functions of the decision variables.

3 Option Pricing with Convex Transaction Costs

In [3], the option pricing problem with bid-ask spread transaction costs is formulated as a linear program. Following the same procedure, we consider the case when the transaction costs are convex. The stock process is discretized into N stages and we use binomial approximation to the stock prices (which can be broadened to include other discrete processes).

For this model, the parameters are the following:

$S(k)$: the stock share price at node k,
$B(k)$: the bond price at node k,
$n(k)$: the quantity of stock shares at node k,
$m(k)$: the quantity of bonds at k,
K: strike price,
$x(k)$: number of shares bought at node k,
$y(k)$: number of shares sold at node k,
k^-: the node that is the immediate predecessor of k,
θ: a parameter reflecting the transaction cost impact.

The option pricing problem can be formulated as the following nonlinear program:

$$\min_{n,m,x,y} \; n(k_0)S(k_0) + m(k_0)B(k_0)$$

subject to

$$\begin{aligned}
& n(k^-)S(k) + m(k^-)B(k) \geq \\
& \qquad n(k)S(k) + m(k)B(k) + \theta((x(k)S(k))^2 + (y(k)S(k))^2) && \text{for every node } k \neq k_0 \\
& n(k) - n(k^-) = x(k) - y(k) && \text{for every node } k \neq k_0 \\
& n(k^-)S(k) + m(k^-)B(k) \geq max(S(k) - K, 0) && \text{for every terminal node } k \\
& x(k) \geq 0, y(k) \geq 0, && \text{for every node } k \neq k_0.
\end{aligned}$$

Again, in this case, higher-order transaction costs lead to more complex nonlinear constraints.

4 Optimization for Index Tracking

The index tracking problem is to find a portfolio of a small number of assets which minimizes a chosen measure of index tracking error or maximizes the correlation between the portfolio and the index portfolio. It is important in these problems to hold relatively few assets. This constraint then prevents a portfolio from holding very small and illiquid positions and limits administration and transaction costs.

4.1 Minimizing Tracking Error

Let x_i represent the percentage weight of asset i in the portfolio x. In [5], Coleman, Li and Henninger formulate the tracking error minimization problem as the following constrained discontinuous optimization problem.

$$\begin{array}{ll} \min_{x \in R^n} TE(x) & \\ \text{subject to} & \sum_{i=1}^{n} \Lambda(x_i) \leq K \\ & \sum_{i=1}^{n} x_i = 1 \\ & x \geq 0, \end{array}$$

where K is the desired number of assets and

$$\Lambda(z) = \begin{cases} 1, & if\ z \neq 0, \\ 0, & otherwise. \end{cases}$$

While the tracking error can be defined in a variety of ways, (particularly with a function that may reward performance above the index), the most common tracking error measure is the following:

$$TE(x) \equiv (x - w)^T Q (x - w), \tag{8}$$

where w denotes the asset weights for the index and Q is the covariance matrix of the asset returns.

Since $\Lambda(x)$ is a discontinuous function, the optimization problem is discontinuous. Coleman, et al., however, use various continuous nonlinear (and nonconvex) approximations of Λ to obtain tractable formulations. Their computational results indicate that close approximations to the discontinuous-constraint solution are possible.

4.2 Maximizing Correlation

This form of the problem can be formulated as follows:

$$\max_x \frac{x^T Q w}{(x^T Q x)^{\frac{1}{2}} (w^T Q w)^{\frac{1}{2}}}$$
$$\text{subject to} \quad \sum_{i=1}^{n} \Lambda(x_i) \le K$$

Since the objective function is nonconvex, the optimization problem is nonconvex. Various approximations are, however, again possible to obtain useful solutions.

5 Optimization in Risk Management

The financial area that has generally received the most quantitative attention is risk management, traditionally for market (i.e., price) risk, and now particularly for credit risk, but increasingly for other forms of risks, such as operational and even model risk. A basic form of risk management optimization is to determine investments x as above with a criterion that the losses with a given probability can be no greater than a certain value. This measure is called *value-at-risk* (VaR), where $P\{L \ge VaR_\alpha\} \le 1 - \alpha$, where L is the loss over a given time interval (day, week, month, year, etc.).

Suppose that prices of the assets are now S_0. Their future value is a random value S_T that depends on a set of risk factors ξ. The loss function is $L(x, S_T) = x^T(S_T - S_0)$. If L then has distribution ψ, $VaR_\alpha = \inf\{l | \psi(s, l) \ge \alpha\}$. We can also define a related quantity, the *conditional value-at-risk* ($CVaR_\alpha$) that has certain properties (*coherence*, see, Rockafellar and Uryasev [7]) that are useful in determining the amount of capital to allocate to cover risks. $CVaR_\alpha$ is defined as:

$$CVaR_\alpha = \inf_l (l + (\frac{1}{1-\alpha}) E[(L(x, S_T(\xi)) - l)^+]), \tag{9}$$

which is the expectation of the losses conditional on their being greater than VaR_α. Now, the problem to minimize $CVaR_\alpha$ for constraints on x becomes a convex nonlinear optimization model, that is reduced to a linear optimization model for discrete realization of ξ. Alexander, Coleman, and Li report in [6] that these problems, however, are ill-posed and lead to significant changes in the x values as prices change. They propose adding a penalty on deviations in x to smooth the solutions.

Alexander, Coleman, and Li also note in [6] that the objective function with m samples, $l + \frac{1}{m(1-\alpha)} \sum_{i=1}^{m} [(x^T(S_0 - S_T(\xi_i)) - l)^+]$, can be approximated by a continuous function for large m. In particular, they suggest a function $p_\epsilon(z)$ for $z = x^T(S_0 - S_T(\xi)) - l$ to approximate z^+ smoothly using a quadratic increasing from $-\epsilon$ to ϵ. This then results in a convex quadratic problem for which they report achieving good performance.

These problems can also be enhanced to include other nonlinearities that track $CVaR$ more closely. In addition, most institutions need to control VaR, which would in general create a nonconvex optimization problem. While $CVaR$ has good properties for optimization, it may not be useful when distributions have unknown tail properties. Very heavy tails may have little effect on VaR but $CVaR$ will change dramatically. For this reason, Heyde and Kou [8], for example, find

VaR to be a more stable measure for risk management. In that case, optimization requires the consideration of greater nonlinearities.

6 Conclusion

Increased sophistication in financial models and consideration of multiple interacting phenomena are leading to greater interest in nonlinear optimization techniques. Solving such problems efficiently can have a direct P&L impact on a variety of firms including trading companies, hedge funds, insurance companies, and traditional banks.

References

[1] T.F. Coleman, Y. Kim, Y. Li, and A. Verman, *Dynamic Hedging in A Volative Market*, Techinal Report, Cornell Theory Center,1999.

[2] T.F. Coleman, Y. Li, and A. Verman, *Reconstructing the Unknown Volatility Function*, Journal of Computational Finance, 2(3): 77-102,1999.

[3] G. Cornuejols and R. Tütüncü, *Optimization Methods in Finance*, Cambridge University Press, 2007.

[4] J. Birge, *Effectively Managing Liquidity Risk in Dynamic Asset-Liability Optimization*, GARP Convention, New York, 2006.

[5] T. F. Coleman, J. Henninger, Y. Li, *Minimizing Tracking Error while Restricting the Number of Assets*, Journal of Risk, vol 8, pp. 33-56, 2006.

[6] S. Alexander, T.F. Coleman, and Y. Li, *Minimizing CVaR and VaR for a portfolio of derivatives,* J. of Banking and Finance 30 (2006), 563-605.

[7] R.T. Rockafellar and S. Uryasev, *Conditional value-at-risk for general loss distributions,* J. of Banking and Finance 26 (2002), 1443-1471.

[8] C.C. Heyde and S.G. Kou, *On the controversy over tailweight of distributions,* Operations Research Letters 32 (2004), 399-408.
