SENIOR RESEARCH PROJECT
Title: Portfolio Optimization and Covariance Matrices
Supervisor: Prof John Mitchell

Controlling risk is one of the primary concerns when allocating money to various financial assets such as stocks, bonds, and money market funds. In order to estimate risk, it is necessary to look at the prior performance of these assets.

In the classical Markowitz portfolio model, expected return is maximized subject to keeping the risk below some acceptable threshold. The expected return is calculated using historical data. Risk is measured using a covariance matrix of historical data. Incorporation of transaction costs [4] into the model and/or considerations of robustness [2] make it difficult to solve portfolio optimization problems.

Another method for deciding upon an allocation of assets is to look at *Value-at-Risk*, or VaR. Finding the allocation that optimizes VaR is a hard optimization problem [1]. To do this accurately, it is necessary to look at historical data. The calculation of VaR depends on the performance of the assets under various scenarios: Given k scenarios, we calculate the return of our portfolio under these scenarios, and then determine a confidence interval. For example, perhaps for 95% of our scenarios the return is at least r%; this value r is then the Value-at-Risk. Given a covariance matrix of historical returns, scenarios can be readily generated.

To be able to investigate different approaches to portfolio optimization it is necessary to collect accurate historical data. Covariance matrices of stock returns do not appear to be readily available. Therefore, it would be useful to construct such a matrix, or collection of matrices [3]. One possible source of data is Yahoo! Finance, which adjusts its historical return data to account for dividends and stock splits.

References

[1] S. Alexander, T. F. Coleman, and Y. Li. Minimizing CVaR and VaR for a portfolio of derivatives. *Journal of Banking and Finance*, 30:583–605, 2006.

[2] E. Erdoğan, D. Goldfarb, and G. Iyengar. Robust portfolio management. Technical Report CORC TR-2004-11, IEOR Department, Columbia University, New York, NY 10027, November 2004.

[3] O. Ledoit and M. Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, 10:603–621, 2003.

[4] J. E. Mitchell and S. Braun. Rebalancing an investment portfolio in the presence of transaction costs. Technical report, Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, November 2002.

SENIOR RESEARCH PROJECT
Title: Approximating LPCCs with SDPs
Supervisor: Prof John Mitchell

My apologies for the jargon in the title!

LPCCs are linear programs with complementarity constraints. Linear programming problems are the fundamental optimization problems. A complementarity constraint requires that at least one of a certain pair of variables should be zero. These problems have applications in finance, data mining, engineering, and elsewhere. An LPCC can be written:

$$\begin{array}{lll} \text{minimize}_{(x,y)} & c^T x + d^T y & \\ \text{subject to} & Ax + By \geq f & \\ \text{and} & 0 \leq y \perp s := q + Nx + My \geq 0, & \end{array} \tag{1}$$

where $a \perp b$ means that the two vectors are orthogonal; i.e., $a^T b = 0$. Further, $c \in \mathbb{R}^n$, $d \in \mathbb{R}^m$, $f \in \mathbb{R}^k$, $q \in \mathbb{R}^m$, $A \in \mathbb{R}^{k \times n}$, $B \in \mathbb{R}^{k \times m}$, $M \in \mathbb{R}^{m \times m}$, and $N \in \mathbb{R}^{m \times n}$ are given. We want to find the optimal $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. For more information see [2].

LPCCs are hard to solve. A lower bound can be found by solving a *relaxation*. Unfortunately, the simple relaxation obtained by dropping the complementarity constraint is typically not very good. This project is concerned with investigating a stronger relaxation which uses a methodology called *semidefinite programming (SDP)*.

We set up the SDP relaxation by introducing new variables, say z_{ij}, that we would like to equal products of variables, of the form $x_i y_j$ etc. We can't directly constrain $z_{ij} = x_i y_j$ because this nonlinear equality constraint is hard to work with. Therefore, we try to approximate the nonlinear equality by linear constraints. For example, by exploiting the fact that $y_i s_i = 0$ for any i, we can place linear constraints on the new variables. Thus,

$$q_i y_i + \sum_{j=1}^{n} N_{ij} x_j y_i + \sum_{j=1}^{m} M_{ij} y_j y_i = 0 \quad \text{for } i = 1, \ldots, m \tag{2}$$

We can also write the new variables as a symmetric square matrix, and we can constrain the matrix to be positive semidefinite. This gives us an SDP relaxation: we have a linear objective function, linear constraints, and some of the variables can be arranged in the form of a matrix which must be positive semidefinite.

This approach has been very successful for a certain class of LPCCs [1]. I am interested in seeing how well it extends to other classes of LPCCs. For example, SDP relaxations of the problems in [2] should be formed and these relaxations can be solved by an SDP package such as CSDP or SDPT3.

References

[1] S. Burer and D. Vandenbussche. Globally solving box-constrained nonconvex quadratic programs with semidefinite-based finite branch-and-bound. Technical report, Department of Management Sciences, University of Iowa, Iowa City, IA, November 2006. *Computational Optimization and Applications*, in print.

[2] J. Hu, J. E. Mitchell, J.-S. Pang, K. P. Bennett, and G. Kunapuli. On the global solution of linear programs with linear complementarity constraints. *SIAM Journal on Optimization*, 19(1):445–471, 2008.