Faculty of Arts and Sciences
Department of Computer Science
CMPS 351 – Numerical Optimization
Assignment 9 – Due Mar 25, 2019

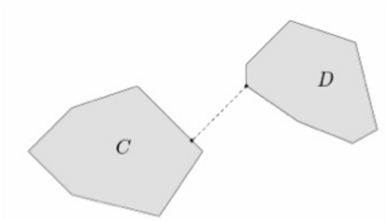
Background reading / reference material:

- Background on portfolio optimization: pp.155-156 and pp.185-187 in reference 2.
- Equality-constrained convex problems: chapter 10 from reference 2.

Announcement:

- Exam will take place on Wednesday April 10 at 6:30pm in Bliss 203.

1. Distance between polyhedra. Consider two polyhedra C and D as shown defined by two sets of linear inequalities $A_i x \leq b_1$ and $A_2 x \leq b_2$. We seek to find the minimum distance between C and D (i.e., the distance between the closest pair of points, one in C and one in D). The problem can be formulated as a quadratic program (quadratic objective, and linear inequalities). In two dimensions, the problem can be formulated in four variables (the coordinates of the closest points). Solve the problem for the data given. Interpret the values obtained for the Lagrange multipliers at optimality.



2. Portfolio Optimization. Consider the problem of allocating an amount of money among a collection of n possible investments held over a period of time whose returns are $r_i = r_1, r_2, \cdots, r_n$. The returns are of course not usually known in advance and are often modeled as random variables that follow normal distributions, each with known expected value and standard deviation μ_i, σ_i.

Investors would like to have the highest return possible from an investment. However, this has to be counterbalanced by the amount of risk the investor is able or desires to take. The expected return and the risk measured by the variance (or the standard deviation, which is the square-root of the variance) are the two main characteristics of a portfolio. Investments with high returns usually correlate with high risk. The mathematical problem of portfolio optimization was initiated by Markowitz in the fifties for which he was rewarded with a Nobel Prize in Economics in 1990.

Let x_i denote the amount of asset i help throughout the period. A normal (also called "long" in investment jargon) position in asset i corresponds to $x_i > 0$. A wide variety of constraints on the portfolio can be considered. The simplest set of constraints is that the sum to be invested is equal to a given budget (usually taken to be 1). Given the stochastic model for the individual asset returns, the total return on the portfolio is defined by $R = \sum x_i \mu_i$ and its variance is $V = \sum_i \sum_j x_i x_j \sigma_i \sigma_j \rho_{ij}$ where ρ_{ij} is the correlation between pairs of returns. This correlation measures the tendency of the return on investments i and j to move in the same direction. Two investments whose return tend to rise and fall together have a positive correlation; the closer ρ_{ij} is to 1, the more closely the two investments track each other. Two investments whose returns tend to move in opposite directions have a negative correlation.

The choice of a portfolio x involves a tradeoff between the mean of the return and its variance. The classical portfolio optimization problem is the quadratic program that seeks to minimize the return variance (which is associated with the risk of the portfolio) subject to achieving a minimum acceptable mean return $r_{\min}$ and satisfying the budget constraint. As a specific example, consider the simple portfolio optimization problem with 4 assets (described on p. 186 of reference 2), with average returns and standard deviations given by the following table:

Asset	1	2	3	4
Avg return	12%	10%	7%	3%
Std Dev	20%	10%	5%	0%

Asset 4 is a risk-free asset, with a (certain) 3% return. Assets 3, 2, and 1 have increasing average returns as well as increasing standard deviations. The correlation coefficients between the assets are: $\rho_{12} = 30\%$, $\rho_{13} = 40\%$, $\rho_{23} = 0\%$.

- Formulate and solve the problem with the data above for a minimum acceptable return of 10%
- The problem can also be formulated as a bi-objective optimization problem, that seeks to maximize return (minimize the negative of the return) and minimize risk (variance). Generate the set of Pareto optimal portfolios that correspond to the data above, and comment on the curve. A Pareto point is generated by combining the 2 objectives into a single scalar objective by means of a parameter κ that may be interpreted as a "risk tolerance parameter".
- An extension to the simple portfolio optimization problem above consists of allowing short positions. A short position in asset i implies the obligation to buy the asset at the end of the investment period and corresponds to $x_i < 0$. Formulate the problem, with non-negative variables only, and add a constraint that the total short position at the beginning of the period is limited to some fraction p of the total long position at the beginning of the period.

3. Newton's Method for equality-constrained convex problems. In this problem you are to explore a solution algorithm that can start at a point x^o that may be infeasible with respect to the equality constraints. This is known as an infeasible start method.

- Implement an infeasible start Newton method with a backtracking line search:

  ```
  def infeasible_newton(f, gf, Hf, A, b, x0)
  ```

  to solve an equality constrained convex problem. Your function should return the primal and dual solution variables (x and ν), and the history of the iterates.
- Use your implementation to solve the following problem starting at $x^o = \mathbf{1}$ and $\nu^o = \mathbf{0}$ ($\mathbf{1}$ and $\mathbf{0}$ are the n-dimensional vectors consisting of ones and zeros, respectively).

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^{n} \log x_i \\ \text{subject to} & Ax = b \end{array} \tag{1}$$

  You may generate random data or use the posted data ($n = 100, p = 40, A_{p\times n}, b_{p\times 1}$) to test your implementation
- Plot the norms of the primal ($\|Ax - b\|$) and dual ($\|\nabla f + A^t \nu\|$) residuals as functions of iteration number. Comment.
- It is possible to perform the linear algebra in a way that avoids factoring the large Hessian of the Lagrangian and instead factor a smaller $p \times p$ matrix $A(\nabla^2 f)^{-1}A^t$, known as a Schur complement. This matrix can be computed efficiently for the problem above. Show how this can be done and how the linear algebra involved in Newton's method may be performed.