FINANCIAL OPTIMIZATION

Lecture 1: General Principles and Analytic Optimization

Philip H. Dybvig
Washington University
Saint Louis, Missouri

A Canonical Optimization Problem

Choose $x \in \Re^N$ to
minimize $f(x)$
subject to $(\forall i \in \mathcal{E}) g_i(x) = 0$, and
$(\forall i \in \mathcal{I}) g_i(x) \geq 0$.

$x = (x_1, ..., x_N)$ is a vector of *choice variables*.
$f(x)$ is the scalar-valued *objective function*.
$g_i(x) = 0$, $i \in \mathcal{E}$ are *equality constraints*.
$g_i(x) \geq 0$, $i \in \mathcal{I}$ are *inequality constraints*.
$\mathcal{E} \cap \mathcal{I} = \emptyset$

- Maximizing $f(x)$ is equivalent to minimizing $-f(x)$.
- Problems in finance are usually written as maximizations.
- Pick the most natural version to communicate your results.
- Sometimes constraints on x are shown separately when specifying the domain of x, e.g., $x \in \Re^N_+$, $x \in [0,1]^N$, or $x \in \{0,1\}^N$.

Solutions

- A *feasible solution* (or *feasible choice*) x satisfies the constraints but may not maximize the objective function. A problem is said to be *feasible* if it has a feasible solution.
- An *optimal solution* (or *optimal choice*) x is feasible and x has the smallest value of the objective function (largest if maximizing) of all feasible solutions. (Also, commonly called "solution" but the book uses this for a candidate that need not be feasible or optimal.)
- An *interior solution* is an optimal solution at which no constraints are binding.
- A *corner solution* is an optimal solution at which constraints are binding.
- A *local optimum* is a feasible choice x^* that, for some $\varepsilon > 0$, is optimal in the problem with the additional constraint $\|x - x^*\| \leq \varepsilon$.
- *value* is $f(x^*)$ for optimal x^*

Drawing Pictures: Solutions

- local optimum
- global optimum
- unbounded problem
- boundary solution
- interior solution

Finding a Sensible Optimization Problem – Example

Find a portfolio strategy for an aggressively managed fund doing market timing. Your colleagues have a model for how expected return evolves as a function of past returns in a binomial model. The reduced form of the model says that there are $N >> 0$ final states of nature and each state of nature n has a price $p_n > 0$ and probability $\pi_n > 0$, where the states have been ordered so that p_n/π_n is strictly increasing in n (the cheapest states are first). Given initial wealth W_0, you should find an optimal strategy that gives the payoff x_n in state n that satisfies the budget constraint $\Sigma_{n=1}^{N} p_n x_n = W_0$.

First Model: In-Class Exercise

Given $p_n > 0$ and $\pi_n > 0$ for $n = 1, ..., N$, and $W_0 > 0$,
choose $x = (x_1, ..., x_N) \in \Re^N$ to
maximize $\Sigma_{n=1}^N \pi_n x_n$
subject to $\Sigma_{n=1}^N p_n x_n = W_0$.

Is this optimization problem feasible?

Does this optimization problem have an optimal solution?

If so, does the optimal solution make sense?

hint: consider changing just two x_is.

Second Model: In-Class Exercise

Given $p_n > 0$ and $\pi_n > 0$ for $n = 1, ..., N$, and $W_0 > 0$,
choose $x = (x_1, ..., x_N) \in \Re^N$ to
maximize $\Sigma_{n=1}^N \pi_n x_n$
subject to $\Sigma_{n=1}^N p_n x_n = W_0$ and
$(\forall n) x_n \geq 0$

Is this optimization problem feasible?

Does this optimization problem have an optimal solution?

If so, does the optimal solution make sense?

hint: look for a corner solution

Kuhn-Tucker Conditions

$\nabla f(x^*) = \Sigma_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \nabla g_i(x^*)$
$(\forall i \in \mathcal{I}) \lambda_i \geq 0$
$\lambda_i g_i(x^*) = 0$

The feasible solution x^* is called *regular* if the set $\{\nabla g_i(x^*) | g_i(x^*) = 0\}$ is a linearly independent set. In particular, an interior solution is always regular.

If x^* is regular and f and the g_is are differentiable, the Kuhn-Tucker conditions are necessary for feasible x^* to be optimal. If the optimization problem is convex (defined by $f(x)$ and all $g_i(x)$'s are convex functions for inequality constraints and affine[1] functions for equality constraints), then the Kuhn-Tucker conditions are sufficient for an optimum. (Recall that a function $q(x)$ is called convex if, for all x_1 and x_2 in the domain of q, and for all $\alpha \in (0, 1)$, $q(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha q(x_1) + (1 - \alpha)x(x_2)$.)

The solution to a convex optimization problem is unique if the objective function is strictly convex.

[1]An *affine* function is linear plus a constant. Some authors define linear as the same as this, others take the constant to be zero in a linear function.

Drawing Pictures: Kuhn-Tucker Conditions

- interior solution
- single constraint binding
- multiple constraints binding
- inflection point
- irregular point

Third Model: In-Class Exercise

Given $p_n > 0$ and $\pi_n > 0$ for $n = 1, ..., N$, and $W_0 > 0$,
choose $x = (x_1, ..., x_N) \in \Re^N_{++}$ to
maximize $\Sigma_{n=1}^N \pi_n \log(x_n)$
subject to $\Sigma_{n=1}^N p_n x_n = W_0$.

Is this optimization problem feasible?

Does this optimization problem have an optimal solution?

If so, does the optimal solution make sense?

hint: solve the Kuhn-Tucker conditions

Shadow Prices

The λ_i's in the objective function are sometimes called *Lagrange multipliers* because of their role in the Lagrangian function that is related to the *Principle of Least Action* in physics, or dual variables because they are related to linear functionals at the solution. More interestingly for us, the λ_i's are shadow prices that measure how much we would pay at the margin to relax the constraint. This interpretation is useful for doing sensitivity analysis to the level of the constraint. For example, if $\lambda_i = 0$, the constraint is not (strictly) binding and relaxing the constraint does not affect the solution. If $\lambda_i = 10$, that means that relaxing the constraint by a small amount $\varepsilon > 0$ should reduce the objective function (in a minimization) by 10ε. This sort of sensitivity analysis is really useful because usually some of the inputs are guesses or policy decisions at another level and we need to know how much they matter.

Disclaimer: this discussion assumes a convex optimization (or nonlocal things might matter) at a regular point (or else the constraint might be redundant).

von Neumann-Morgenstern Preferences

Choosing to maximize the expectation of $\log(x)$ is a special case of von Neumann-Morgenstern preferences. The general form of these preferences maximizes the expectation of $u(x)$ where $u(\cdot)$ is the von Neumman-Morgenstern utility function. These preferences were given an axiomatic foundation by John von Neumann and Oskar Morgenstern, and are the most commonly-used preferences in financial research. The most popular preference assumption in practice is mean-variance utility, which has a lot of conceptual problems but is easy to work with. Academic researchers are looking at many other preference assumptions that go beyond these traditional models by including such features as preference for the timing of resolution of uncertainty and ambiguity aversion. These topics in Choice Theory are interesting but beyond the scope of this course.

Convex Optimization: Intuition

The Kuhn-Tucker conditions are local conditions that only look at the value and derivative of the objective and contraint functions at a point. For a convex optimization problem, this is good enough because any local solution is a global solution.

For a convex optimization problem, the feasible set is convex, that is, for all feasible x_1 and x_2 and for all $\alpha \in (0, 1)$, $\alpha x_1 + (1 - \alpha)x_2$ is also feasible. (This follows easily from the definitions.) With convexity of the objective function, this implies that if another feasible choice is better, so are many nearby feasible choices. (This also follows easily from the definitions.) Therefore, any point that is not a global optimum is not a local optimum, which is equivalent to saying that any point that is a local optimum is also a global optimum.

Drawing Pictures: Convex and Nonconvex Optimization

- local optimum = global optimum?
- typical algorithm: gradient and Newton-Raphson

Why Convexity Matters in Practice

With some interesting exceptions (such as integer linear programs and certain continuous-time dynamic problems), choice problems we can solve reliably are convex optimization problems. These problems tend to be easier to solve reliably because a choice is optimal if and only if it is optimal locally (within some neighborhood) and the algorithm is not going to be confused by a local optimum. Also, optimization algorithms have a clear path through the feasible set to the optimum and globally good local indications about how to make way towards an optimum.

A sufficiently smooth function of one variable is convex if its second derivative is everywhere nonnegative. A sufficiently smooth real-valued function of many variables is convex if its matrix of second derivatives is everywhere positive semidefinite (all eigenvalues are nonnegative). If your choice problem is convex, you have some reason to expect your nonlinear optimization problem will converge on an optimal solution, and if it is not convex, you have every reason to think numerical optimization might have problems and maybe you should double-check the solution..

Mean-variance Analysis: First Pass

Choose portfolio weights x_1, ..., x_N to
maximize $x_0 r_F + \left(\Sigma_{i=1}^N x_i \mu_i\right) - (1/2)\gamma \left(\Sigma_{i=1}^N \Sigma_{j=1}^N x_i \sigma_{ij} x_j\right)$, subject to
$\Sigma_{i=0}^N x_i = 1$.

$\mu = (\mu_i)$: vector of means
$\Sigma = (\sigma_{ij})$: covariance matrix, positive definite
$\gamma > 0$: risk aversion parameter

Mean-variance Analysis: Solution and Observation

First-order (Kuhn-Tucker) condition:

$$\mu_0 = \lambda$$

$$\mu_i = \gamma \sum_{j=1}^{N} \sigma_{ij} x_j + \lambda \quad \text{for } i > 0$$

or in vector notation (using $\mathbf{1}$ to denote a vector of 1s of length N):

$$\mu = r_F \mathbf{1} + \gamma \Sigma^\top x$$

which implies

$$x = \frac{1}{\gamma} \Sigma^{-1} (\mu - r_F \mathbf{1})$$

This is the most common financial application of optimization, but it is rarely done competently because people do not use reasonable choices of μ and Σ (sample estimates are dominated by estimation error) and compensate by adding *ad hoc* constraints that help but do not fix the problem.

Summary

- Optimization problems have choice variables, objective functions and constraints
- Optimal solutions and feasible solutions, interior solutions and corner solutions, local solutions
- Kuhn-Tucker conditions
- Shadow prices are useful for sensitivity analysis
- Convex optimization is safest