Lecture 5

Model and Solutions for Convex Optimization Problems

Qiong Wang

Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

September 18-20, 2023

Fortune's Formula

- 1948: Claude Shannon (Information Theory) "The Mathematical Theory of Communication".
- 1956: John Kelly (Kelly Criterion), "A New Interpretation of Information Rate".
- 1966: Edward Thorp, "Beat the Dealer: A Winning Strategy for the Game of Twenty-One".
- 1967: Edward Thorp and Sheen Kassouf, "Beat the Market: A Scientific Stock Market System".
- 1999: Edward Thorp, "The Kelly Criterion in Blackjack Sports Betting and the Stock Market".
- 2006: William Poundstone, "Fortune's Formula: The Untold Story of the Scientific Betting System That Beat the Casinos and Wall Street".

Kelly's Criterion:Simple Case

Problem Statement: Suppose that you can make repeated investments on something that generates $(1+b)$ for each dollar invested if you succeed and results a complete loss (0) if you fail. The probability of success is p. Starting from a fixed initial budget, you can keep investing as long as you have money left. Your objective is to maximize your wealth after N^{th} round, where N is a large number. How much should you invest in each round?

Model Formulation: let x be % of budget that you invest

	amount not invested	return	total
win	$1-x$	$x(1+b)$	$1+bx$
lose	$1-x$	0	$1-x$

expected wealth after N^{th} round (K: number of times you win)

$$W_N(x) = (1+bx)^K(1-x)^{N-K}.$$

Question: how to choose x to maximize $W_N(x)$.

Problem Transformation

$$\max_{0\le x\le 1}\{W_N(x)\} \quad \text{where} \quad W_N(x) = (1+bx)^K(1-x)^{N-K}.$$

- how does the optimal solution depend on N (besides that it is large)?
- how to determine the value of K?

take log

$$g_N(x) = \frac{\ln W_N(x)}{N} = \frac{\ln(1+bx)^K + \ln(1-x)^{N-K}}{N}$$
$$= \frac{K}{N}\ln(1+bx) + \frac{N-K}{N}\ln(1-x)$$

1. for any given N, maximizing $W_N(x)$ is the same as maximizing $g_N(x)$.
2. by the Law of Large Numbers

$$\boxed{\lim_{N\to\infty}\frac{K}{N} = p,}$$

3. so as N gets large, $g_N(x)$ converges to

$$g(x) = p\ln(1+bx) + (1-p)\ln(1-x)$$

Optimal Solution of Kelly's Criterion (Simple Case)

$$\max_{0 \le x \le 1} \{p \ln(1+bx) + (1-p)\ln(1-x)\}.$$

- take derivative and set it to zero:

$$p\frac{b}{1+bx} - (1-p)\frac{1}{1-x} = 0.$$

dont forget constraints *

solve the equation

$$\frac{1+bx}{pb} = \frac{1-x}{1-p} \quad \longrightarrow \quad x = p - \frac{1-p}{b}.$$

Is that all?.

$$\boxed{x^* = \max\left(p - \frac{1-p}{b}, 0\right)}.$$ *

- when is it optimal to invest all $(x^* = 1)$? if and only if $p = 1$.
- when is it optimal to invest no money $(x^* = 0)$? if and only if $pb \le (1-p)$.

Example: Kelly's Criterion

The General Case:

Suppose that you can make repeated investments on a set of S scenarios. One and only one scenario realizes and the probability of getting scenario i is p_i $(1 \leq i \leq S)$, so

$$p_1 + ... + p_S = 1.$$

If scenario i realizes, you receive $1 + b_i$ $(1 \leq i \leq S)$ from each dollar bet on that scenario and lose all your bets on other scenarios.

Question: What percentage of your budget should you bet on each scenario?

Problem Formulation: Decisions and Objective

decisions: x_0: % of fund to keep on hand,
x_i: % of fund to bet on scenario i ($i = 1, ..., S$).

- N bets, number of times of having scenario i is K_i ($i = 1, ..., S$),
$$K_1 + ... + K_S = N.$$
- if scenario i becomes true: your wealth is multiplied by a factor of
$$x_0 + (1 + b_i)x_i, \qquad i = 1, ..., S.$$
- the objective is to maximize
$$[x_0 + (1 + b_1)x_1]^{K_1} \times ..... \times [x_0 + (1 + b_S)x_S]^{K_S}.$$
- objective function transformation: take log, divide by N,
$$\frac{K_1}{N} \ln[x_0 + (1 + b_1)x_1] + ... + \frac{K_S}{N} \ln[x_0 + (1 + b_S)x_S]$$
apply the Law of Large Numbers:
$$p_1 \ln[x_0 + (1 + b_1)x_1] + ... + p_S \ln[x_0 + (1 + b_S)x_S].$$

Problem Formulation: Constraints

objective: $$\sum_{i=1}^{S} p_i \ln[x_0 + (1 + b_i)x_i]$$ maximize

constraints: $$x_0 \geq 0, x_1 \geq 0, ..., x_S \geq 0$$

$$x_0 + ... + x_S = 1 \longrightarrow \begin{cases} x_0 + ... + x_S \leq 1, \\ x_0 + ... + x_S \geq 1. \end{cases}$$

relax to

easier to convert to 2 constraints

question: do we really need the constraint $x_0 + ... + x_S \geq 1$?

answer: no (why?)

problem completely defined by:

$$\max_{x_0,\ldots,x_S} \left\{ \sum_{i=1}^{S} p_i \ln[x_0 + (1 + b_i)x_i] \right\}$$

subject to
$$x_0 + ... + x_S \leq 1,$$
$$x_0 \geq 0, x_1 \geq 0, ..., x_S \geq 0.$$

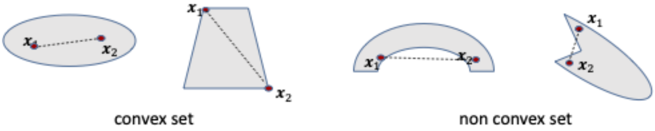
Convex Optimization Problem

$$\max_{\mathbf{x}} \{g(\mathbf{x})\} \quad \text{subject to} \quad h_j(\mathbf{x}) \leq 0, \quad j = 1, .., J.$$

- $g(\mathbf{x})$ is a concave function of $\mathbf{x}$ (convex if min. $g(x)$, i.e., max. $-g(x)$)
- the set of all feasible solutions is convex (i.e., $h_j(x)$ are convex functions):
  - for all $j = 1, ..., J$, $h_j(\mathbf{x})$ is a convex function ($0 \leq \alpha \leq 1$):

$$x = \alpha x_1 + (1-\alpha)x_2 \quad \longrightarrow \quad h_j(x) \leq \alpha h_j(\mathbf{x}_1) + (1-\alpha)h_j(\mathbf{x}_2).$$

  - if $\mathbf{x}_1$ and $\mathbf{x}_2$ are feasible, $\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$ is feasible ($0 \leq \alpha \leq 1$).



Implications of Concave Objective Function

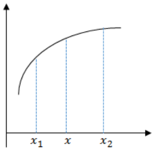
$g(x)$ is a concave function: any x_1, x_2, and $0 \leq \alpha \leq 1$:

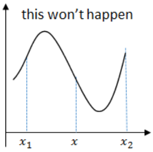
$$x = \alpha x_1 + (1-\alpha)x_2 \quad \longrightarrow \quad \alpha g(x_1) + (1-\alpha)g(x_2) \leq g(x)$$

$$\text{so:} \quad \alpha g(x_1) + (1-\alpha)g(x_2) \leq \alpha g(x) + (1-\alpha)g(x),$$

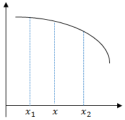
$$\text{i.e.} \quad (1-\alpha)(g(x_2) - g(x)) \leq \alpha(g(x) - g(x_1)).$$

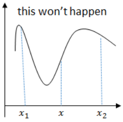
if $g(x_2) \geq g(x)$, then $g(x) \geq g(x_1)$,





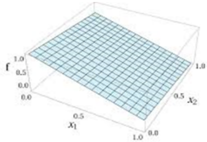
if $g(x) \leq g(x_1)$, then $g(x_2) \leq g(x)$.





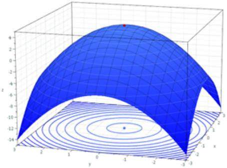
How to Solve a Convex Optimization Problem

- setting the derivative to zero?



what if the optimal solution is at the boundary?

- evaluate solutions at boundaries of the feasible region (like we do with LP)?



what if the optimal solution is in the interior of the feasible region?

Karush-Kuhn-Tucker (KKT) Method

$$\max_{\mathbf{x}} \{g(\mathbf{x})\} \quad \text{subject to} \quad h_j(\mathbf{x}) \leq 0, \quad j = 1, .., J.$$

$g(\mathbf{x})$ is concave and $h_j(\mathbf{x})$ $(j = 1, ..., J)$ are convex functions.

- attach a penalty λ_j, to each constraint $h_j(\mathbf{x}) \leq 0$, $j = 1, ..., J$
- choose $\mathbf{x}$ to maximize

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = g(\mathbf{x}) - \sum_{j=1}^{J} \lambda_j h_j(\mathbf{x}),$$

by setting their derivatives to zero.
(definition: λ_j are Lagrange multipliers; $\mathcal{L}$ is the Lagrangian).

- impose following conditions on the solution: for all $j = 1, ...J$,

$$\begin{aligned} \lambda_j &\geq 0, \\ h_j(\mathbf{x}) &\leq 0, \\ \lambda_j h_j(\mathbf{x}) &= 0. \end{aligned}$$

Optimal Solution

let $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ be a solution of

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial x_i} = \frac{\partial g}{\partial x_i} - \sum_{j=1}^{J} \lambda_j \frac{\partial h_j}{\partial x_i} &= 0, \quad i = 1, ..., n, \\
\lambda_j h_j(\mathbf{x}) &= 0, \quad j = 1, ..., J, \\
\lambda_j &\geq 0, \quad j = 1, ..., J, \\
h_j(\mathbf{x}) &\leq 0, \quad j = 1, ..., J.
\end{aligned}$$

- given $\boldsymbol{\lambda}^*$, because $\partial \mathcal{L}/\partial x_i = 0$ for all $i = 1, \cdots, m$, $\mathbf{x}^*$ maximizes

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*) = g(\mathbf{x}) - \sum_{j=1}^{J} \lambda_j^* h_j(\mathbf{x}), \quad (\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*) \text{ is concave in } \mathbf{x}).$$

- $h_j(\mathbf{x}^*) \leq 0$ for all j, so $\mathbf{x}^*$ is feasible.
- to show $\mathbf{x}^*$ maximizes $g(\mathbf{x})$: let $\mathbf{x}$ be a feasible solution:

$$g(\mathbf{x}) \leq g(\mathbf{x}) - \sum_{j=1}^{J} \lambda_j^* h_j(\mathbf{x}) = \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}^*) \leq \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = g(\mathbf{x}^*).$$

(because $\lambda_j^* h_j(\mathbf{x}^*) = 0$, $j = 1, \cdots, J$).

Summary on KKT Conditions

for convex optimization problem (with a concave objective function to *maximize* and a convex feasible region):

- transform "hard" constraints into "soft" penalties.
- Penalty (aka Lagrange multipliers) for violating the constraint should be
  1. positive;
  2. sufficiently high to prevent any violation of restrictions;
  3. zero if constraints are strictly satisfied (aka complementary slackness).
- the optimal solution is characterized by the first-order conditions for maximizing the Lagrangian and the above three conditions on the penalty.

Back to Kelly's Criterion (the general case)

$$\max_{x_0,\ldots,x_S} \left\{ \sum_{i=1}^{S} p_i \ln[x_0 + (1+b_i)x_i] \right\}$$

subject to

$$x_0 + x_1 + \ldots + x_S \leq 1,$$
$$x_i \geq 0, \quad i = 1, \ldots, S.$$

$i = 1, \ldots, S$ scenarios;

p_i: probability of scenario i;

b_i: return in scenario i;

x_i: % of budget to bet on scenario i;

x_0: cash to keep on hand.

Lagrangian:

$$\mathcal{L}(\mathbf{x}, \lambda) = \sum_{i=1}^{S} p_i \ln[x_0 + (1+b_i)x_i] - \bar{\lambda}(x_0 + \ldots. + x_S - 1) + \sum_{i=0}^{S} \lambda_i x_i,$$

λ_i: Lagrange multiplier attached to $x_i \geq 0$ $(i = 0, \ldots, S)$.

$\bar{\lambda}$: Lagrange multiplier attached to $x_0 + \ldots + x_S \leq 1$.

KKT Conditions

$$\mathcal{L}(\mathbf{x}, \lambda) = \sum_{i=1}^{S} p_i \ln[x_0 + (1+b_i)x_i] - \bar{\lambda}(x_0 + ... + x_S - 1) + \sum_{i=0}^{S} \lambda_i x_i$$

- derivatives of $\mathbf{x}_i$ $(i = 1, \cdots, S)$ to zero:
$$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{p_i(1+b_i)}{x_0 + (1+b_i)x_i} - \bar{\lambda} + \lambda_i = 0, \quad i = 1, .., S,$$

- derivatives of $\mathbf{x}_0$ to zero:
$$\frac{\partial \mathcal{L}}{\partial x_0} = \sum_{i=1}^{S} \frac{p_i}{x_0 + (1+b_i)x_i} - \bar{\lambda} + \lambda_0 = 0.$$

- constraints: $x_0 + x_1 + ... + x_S \leq 1$ and $x_i \geq 0, \quad i = 0, ..., S.$
- complementary slackness condition:
$$\bar{\lambda}(x_0 + x_1 + ... + x_S - 1) = 0 \;\text{ and }\; \lambda_i x_i = 0, \quad i = 0, ..., S.$$
- nonnegative penalties: $\bar{\lambda} \geq 0$ and $\lambda_i \geq 0, \quad i = 0, ...S.$

Investment Insight I:

Observation: from the condition that

$$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{p_i(1+b_i)}{x_0 + (1+b_i)x_i} - \bar{\lambda} + \lambda_i = 0,$$

can this equality hold if we let both $x_0 = 0$ and $x_i = 0$?

- what investment strategy corresponds to $x_0 = 0$ and $x_i = 0$?
- under this strategy, what happens if scenario i realizes?
- so to satisfy

  $$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{p_i(1+b_i)}{x_0 + (1+b_i)x_i} - \bar{\lambda} + \lambda_i = 0 \text{ for all } \quad i = 1, \cdots, S,$$

  what types of solutions can we have?
- correspondingly, what does the solution tell us how to invest?

Investment Insight II: a little algebra first

- for x_i to be optimal:

$$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{p_i(1+b_i)}{x_0+(1+b_i)x_i} - \bar{\lambda} + \lambda_i = 0, \quad i = 1,..,S,$$

so

$$\frac{p_i}{x_0+(1+b_i)x_i} = \frac{\bar{\lambda}-\lambda_i}{1+b_i}, \quad i = 1,...,S. \tag{1}$$

- for x_0 to be optimal:

$$\frac{\partial \mathcal{L}}{\partial x_0} = \sum_{i=1}^{S} \frac{p_i}{x_0+(1+b_i)x_i} - \bar{\lambda} + \lambda_0 = 0 \tag{2}$$

- use (**??**) to replace $p_i/(x_0+(1+b_i)x_i)$ in (**??**):

$$\sum_{i=1}^{S} \frac{\bar{\lambda}-\lambda_i}{1+b_i} - \bar{\lambda} + \lambda_0 = 0$$

i.e.,

$$\bar{\lambda}\left(\sum_{i=1}^{S}\frac{1}{1+b_i} - 1\right) = -\lambda_0 + \sum_{i=1}^{S}\frac{\lambda_i}{1+b_i}.$$

Investment Insight II: observation

- repeat: first-order condition of x_0:

$$\sum_{i=1}^{S} \frac{p_i}{x_0 + (1+b_i)x_i} - \bar{\lambda} + \lambda_0 = 0.$$

implication: $\bar{\lambda} > 0$.

- repeat: from the last slide,

$$\bar{\lambda}\left(\sum_{i=1}^{S} \frac{1}{1+b_i} - 1\right) = -\lambda_0 + \sum_{i=1}^{S} \frac{\lambda_i}{1+b_i}.$$

since $\lambda_i \geq 0$ for all $i = 1, \cdots, S$, if

$$\text{if} \quad \sum_{i=1}^{S} \frac{1}{1+b_i} < 1, \text{then } \lambda_0 > 0$$

so $x_0 = 0$ (because $\lambda_0 x_0 = 0$) and $x_i > 0$ for all $i = 1, \cdots S$.

why?

Investment Insight II: Explanation

Example: $S = 2$, $p_1 = p_2 = 0.5$ (i.e., two scenarios with equal chance), and $b_1 = b_2 = 2$ (i.e., i.e., 200% profit for the correct bet).

$$\frac{1}{1+b_1} + \frac{1}{1+b_2} = \frac{2}{3} < 1,$$

do you want to keep some cash on hand in this situation?

what about any other values of p_1 and p_2?

More generally, suppose that you let

$$x_i = \frac{1/(1+b_i)}{1/(1+b_1) + ... + 1/(1+b_S)}, \quad i = 1, ..., S$$

if scenario i occurs,

$$x_i(1+b_i) = \frac{1}{1/(1+b_1) + ... + 1/(1+b_S)}$$

so in any scenario, your return:

$$x_i(1+b_i) > 1 \ \text{ if } \ \sum_{i=1}^{S} \frac{1}{1+b_i} < 1.$$

Investment Insight II: Discussion

What happened: we infer from the KKT condition that when

$$\sum_{i=1}^{S} \frac{1}{1+b_i} < 1,$$

you should bet everything because there is free profit.

Arbitrage!

remember Fundamental Theorem of Asset Pricing? does it apply here?

- each scenario is an asset, current price 1, future value $1 + b_i$ or 0.
- what is the risk-neutral probability of scenario i ($i = 1, \cdots, S$)?

$$1 = \hat{p}_i(1 + b_i) \quad \longrightarrow \quad \hat{p}_i = \frac{1}{1+b_i}, \ i = 1, .., S.$$

- can such probabilities exist if:

$$\sum_{i=1}^{S} \frac{1}{1+b_i} < 1.$$

Investment Insight: Optimization

$$\text{when} \quad \sum_{i=1}^{S} \frac{1}{1+b_i} < 1:$$

- arbitrage opportunity, bet every penny: $x_0 = 0$, and $x_i > 0$ ($i = 1, ..., S$).
- since $\lambda_i x_i = 0$, $\lambda_1 = ... = \lambda_S = 0 \quad (i = 1, ..., S)$.
- apply $x_0 = 0$ and $\lambda_i = 0$ ($i = 1, \cdots, S$) to the first-order condition:

$$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{p_i(1+b_i)}{x_0 + (1+b_i)x_i} - \bar{\lambda} + \lambda_i = 0, \quad i = 1, ..., S,$$

$$\frac{p_1}{x_1} = ... = \frac{p_S}{x_S} = \bar{\lambda} \qquad \text{and} \qquad x_1 + ... + x_S = 1.$$

- so the optimal solution: $x_i^* = p_i$ ($i = 1, \cdots, S$). Make sense?