Session 11

Revisit the life insurance question.

$$P(Y=1) = p = \frac{\exp(a^T x + b)}{1 + \exp(a^T x + b)}$$

optimization variable x.

constraints $\boxed{Fx \leq g}$

Done: $\max_x P(Y=1)$

Formulation of above problem as an optimization.

Question 2: Maximizing expected profit.

Consider $c^T x + d$ is the profit, and it is positive.

Goal: Maximize the expected profit

Expect profit = probability × value

$$\Rightarrow \underline{E(c^T x + d)} = \frac{\exp(a^T x + b)}{1 + \exp(a^T x + b)} (c^T x + d)$$

(notation)

Let us define, $h(u) = \dfrac{e^u}{1+e^u}$, $u = a^T x + b$

Problem can be written as (see the goal in the last page)

$$\max_x \; \underline{h(a^T x + b)} \, (c^T x + d)$$

constraint $Fx \leq g$

<u>Challenge</u>: the objective is not convex.

(<u>check</u>: $h(a^T x + b)$ is not convex function

$c^T x + d$ is an affine function)

Define a new variable $z = \log(\underline{1 + \exp(a^T x + b)})$

Then $h(u) = \dfrac{e^u}{1+e^u} = 1 - \dfrac{1}{1+e^u}$.

Take $u = a^T x + b$, then, $\boxed{h(u) = 1 - e^{-z}}$

(check it)

The objective function becomes

$$(1 - e^{-z})(c^T x + d) \quad \text{(check)}$$

As $(c^T x + d)$ is ≥ 0 and as $(1-e^{-z})$ is concave function of z (verify it).

Therefore, the objective function is product - affine function and a concave function.

That implies that the objective function is concave function

This is not true in general (see next page)

(Verify that product of affine function & concave function is also concave?)

Thus the problem can be formulated as

$$\max \ (1-e^{-z})(c^T x + d)$$

$$\text{subject to } \begin{cases} z = \log\left(1 + \underline{\exp(a^T x + b)}\right) \\ \underline{Fx \leq g} \end{cases}$$

Study and come back.

Post script (Correction to the maximizing the expected gain)

Recall the objective function:

$$\max_{x} \left(1-e^{-z}\right)\left(c^T x + d\right)$$

subject to $z = \log\left(1+\exp(a^T x + b)\right)$

& $F^T x \leq g$

Note that, the first term of the objective function $(1-e^{-z})$ is a log-concave function (check it)

The second term of the objective function is $c^T x + d$ is positive affine function. It can be shown that $c^T x + d$ is log concave (check by taking double derivative of $\log(a^T x + b)$).

The pointwise product of two log-concave function is also log concave (see Section 3.5.2, page 105, Boyd's book)

The optimization problem becomes optimizing a log-concave function.

$$\max_{x} \left(1-e^{-z}\right)\left(c^T x + d\right)$$

subject to $z = \log\left(1+\exp(a^T x + b)\right)$

$F^T x \leq g$

Lagrangian dual function (Chapter 5, page 215)

<u>Section 5.1.1</u> (of text book).

Consider an optimization in the standard form

$$\left.\begin{aligned} &\text{minimize } f_0(x) \\ &\text{subject to } f_i(x) \le 0, \quad i=1,\dots,m \\ &\qquad\qquad\quad h_i(x) = 0, \quad i=1,\dots,p \end{aligned}\right\} - (5.1) \text{ of the book.}$$

optimization variable $x \in \mathbb{R}^n$.

We assume its domain $D = \bigcap_{i=1}^{m} \mathrm{dom} f_i \bigcap_{i=1}^{p} \mathrm{dom} h_i$ is non empty, and denote the optimal value of (5.1) as p^*

$$\boxed{\text{Note that we don't assume that convexity of (5.1)}}$$

Need to <u>check that</u> (5.1) is a general case of convex optimization.

(Do it later).

Basic idea of Lagrangian duality is that augmenting the objective function with weighted sum of constraint function.

$\boxed{\text{Lagrangian}}$ $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ associated with (5.1) as

$$L(x, \underline{\lambda}, \underline{\nu}) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

where
dom L = dom $f_0 \times \mathbb{R}^m \times \mathbb{R}^p$

λ_i : Lagrangian multiplier with $f_i(x) \leq 0$

ν_i : " " with $h_i(x) = 0$

* Vectors λ and ν are called the dual variables or Lagrange multiplier vectors.

Dual function

$$\underline{g(\lambda,\nu)} = \inf_{x \in D} L(x, \lambda, \nu)$$

Note for any $\lambda \geq 0$, and any $\nu \in \mathbb{R}$,

$$\boxed{g(\lambda,\nu) \leq p^*}$$ (Prove it $\underline{\underline{5.1.3}}$)

$\Rightarrow$ $g(\lambda,\nu)$ provides the bounds of optimal $\underline{p}^*$.

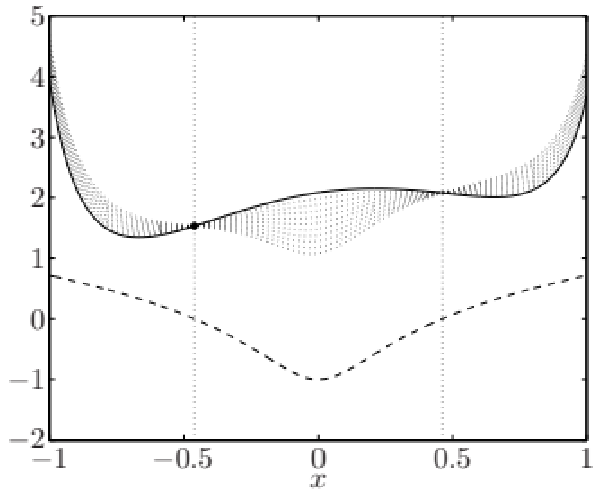

Figure 5.1 *Lower bound from a dual feasible point.* The solid curve shows the objective function f_0, and the dashed curve shows the constraint function f_1. The feasible set is the interval $[-0.46, 0.46]$, which is indicated by the two dotted vertical lines. The optimal point and value are $x^\star = -0.46$, $p^\star = 1.54$ (shown as a circle). The dotted curves show $L(x, \lambda)$ for $\lambda = 0.1,\ 0.2, \ldots, 1.0$. Each of these has a minimum value smaller than $p^\star$, since on the feasible set (and for $\lambda \geq 0$) we have $L(x, \lambda) \leq f_0(x)$.