Minimization problem on p. 113

$$\underset{\underline{w},b}{\text{minimize}} \quad \frac{\|\underline{w}\|^2}{2} \quad \text{s.t.} \quad t_\ell(\underline{w}^T \underline{y}_\ell + b) - 1 \geq 0$$

$$g_\ell(\underline{w}, b) \qquad \text{for } \ell = 1, \dots, L$$

Following Boyd's and Lindenberghe's book p. 215:

Again change from $\underline{w}, b$ to $\underline{x}$.

Formulate constrained minimization problem as

$$\boxed{\min_{\underline{x} \in D} f_0(\underline{x}) \quad \text{s.t.} \quad f_i(\underline{x}) \leq 0 \quad , \; i = 1, \dots, m \quad (P) \quad \text{(disregard equality constraints)}}$$

Is it possible to convert (P) into a convex optimization problem?

Consider the Lagrangian

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

$\underline{x}$: n-dimensional, $\underline{\lambda}$: m-dimensional

with $\underline{\lambda} = [\lambda_1, \dots, \lambda_m]^T$

Assume $\underline{x}$ to be feasible. Then

$$\sup_{\boxed{\underline{\lambda} \geq \underline{0}}} L(\underline{x}, \underline{\lambda}) = \sup_{\underline{\lambda} \geq \underline{0}} \left[f_0(\underline{x}) + \sum_{i=1}^{m} \lambda_i f_i(\underline{x}) \right]$$

$$= \begin{cases} f_0(\underline{x}) & \text{if } f_i(\underline{x}) \leq 0 \text{ for } i = 1, \dots, m \quad \text{CASE A: feasible } \underline{x} \\ \infty & \text{otherwise (CASE B)} \end{cases}$$

CASE A: choose $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$

CASE B: choose $\lambda_j \to \infty$ for $f_j(\underline{x}) > 0$

and $\lambda_k = 0$ for $f_k(\underline{x}) \leq 0$

$\Rightarrow$ Forget CASE B? Just concentrate on case A?

$\Rightarrow$ $\boxed{\text{Thus, for feasible } \underline{x}\text{, we can reinterpret the objective function } f_0(\underline{x}) \text{ as } \sup_{\underline{\lambda} \geq \underline{0}} L(\underline{x}, \underline{\lambda})}$

$\Rightarrow$ Thus, the solution of (P) (called p^*) is given by

$$p^* = \inf_{\underline{x} \in \underline{D}} f_0(\underline{x}) \qquad \text{s.t. constraints}$$

$$= \inf_{\underline{x} \in D} \sup_{\underline{\lambda} \geq \underline{0}} L(\underline{x}, \underline{\lambda}) \quad \text{s.t. constraints}$$

THAT is the main idea of using the Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$

<u>Concept of the DUAL function</u>

Assume that p^* is a solution to the minimization problem

$$p^* = \min_{\underline{x}} f_0(\underline{x}) \quad \text{s.t. } f_i(\underline{x}) \leq 0 \text{ for } i=1,\dots,m$$

Define the <u>dual</u> function

$$g(\underline{\lambda}) = \inf_{\underline{x} \in D} L(\underline{x}, \underline{\lambda}) = \inf_{\underline{x} \in D} \left[f_0(\underline{x}) + \sum_{i=1}^{m} \lambda_i f_i(\underline{x}) \right]$$

for a general vector $\underline{\lambda} \in \mathbb{R}^m$.

If $L(\underline{x}, \underline{\lambda})$ is unbounded below in $\underline{x}$, we have $g(\underline{\lambda}) = -\infty$. Note that in the reformulation of the minimization problem above, we have in the Lagrangian that $\underline{\lambda} \geq \underline{0}$ and

$$p^* = \inf_{\underline{x} \in D} f_0(\underline{x}) = \inf_{\underline{x} \in D} \sup_{\lambda \geq 0} L(\underline{x}, \underline{\lambda})$$

How can we relate $g(\underline{\lambda}) = \inf_{\underline{x} \in D} L(\underline{x}, \lambda)$ to

$p^* = \inf_{\underline{x} \in D} \sup_{\underline{\lambda} \geq \underline{0}} L(\underline{x}, \underline{\lambda})$?

It turns out that for $\underline{\lambda} \geq \underline{0}$

$$\boxed{g(\underline{\lambda}) \leq p^* \leq f_0(\underline{x}). \qquad (*)}$$

Proof: Assume $\tilde{\underline{x}}$ to be a feasible point, i.e. $f_i(\tilde{\underline{x}}) \leq 0$ for $i = 1, \dots, m$ and $\underline{\lambda} \geq \underline{0}$

$$\Rightarrow \sum_{i=1}^{m} \underbrace{\lambda_i}_{\text{non-negative}} \underbrace{f_i(\tilde{\underline{x}})}_{\text{non-positive}} \leq 0$$

Clearly, from definition of $g(\underline{\lambda})$

$$g(\underline{\lambda}) = \inf_{x \in D} L(\underline{x}, \underline{\lambda}) \leq L(\tilde{\underline{x}}, \underline{\lambda}) \qquad (**)$$

and, since $\tilde{x}$ is feasible, we have

$$L(\tilde{\underline{x}}, \underline{\lambda}) \leq f_0(\tilde{\underline{x}}) \qquad (***)$$

From $(**)$ and $(***)$, it follows that

$$g(\underline{\lambda}) \leq f_0(\tilde{\underline{x}}) \qquad (****)$$

Finally, since $(****)$ holds for any feasible point (with $\underline{\lambda} \geq \underline{0}$), so it does for p^*.

In particular, we can conclude that

$$\underbrace{d^*}_{\text{dual solution}} = \sup_{\underline{\lambda} \geq \underline{0}} g(\underline{\lambda}) = \sup_{\lambda \geq 0} \inf_{\underline{x} \in D} L(\underline{x}, \underline{\lambda})$$

$$\leq p^* = \inf_{\underline{x} \in D} \sup_{\lambda \geq \underline{0}} L(\underline{x}, \underline{\lambda})$$

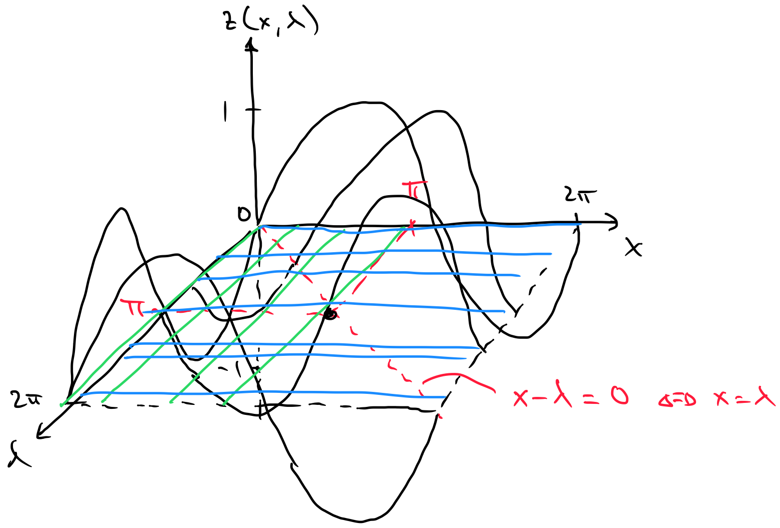
"saddle point theorem"

Example: function $z(x, \lambda) = \sin(x - \lambda)$ with $(x,\lambda) \in [0,2\pi)^2$

(x, λ: scalars)

(explain the saddle point, namely,

$\sup \inf \leq \inf \sup$)



$$\inf_{x \in [0,2\pi)} \sup_{\lambda \geq 0} z(x,\lambda) = \inf_{x \in [0,2\pi)} 1 = 1 = p^*$$

$$\sup_{\lambda \geq 0} \inf_{x \in [0,2\pi)} z(x,\lambda) = \sup_{\lambda \geq 0} (-1) = -1 = d^*$$