- Primal problem v.s. Dual problem

$\min f_0(x)$ $\max g(\lambda,\nu)$

s.t.

$f_i(x)\le 0$ $\lambda\ge 0$

$h_i(x)=0$

$$\min_{x\in D}\max_{\lambda,\nu} L(x,\lambda,\nu) = \max_{\lambda,\nu}\min_{x\in D} L(x,\lambda,\nu)$$

minimax theorem

- Minimax property:

ex:

		z		
w	1	-1	3	$\to$ 3
	2	2	-1	$\to$ 2
	3	1	-2	$\to$ 3
	$\downarrow$ 1	$\downarrow$ -1	$\downarrow$ -2	

$\Rightarrow$ ① maximize then ② minimize seems to induce large value

Proof: for arbitrary $\tilde{w}, \tilde{z}$

$$\min_{w\in W} f(w,\tilde{z}) \le f(\tilde{w},\tilde{z}) \le \max_{z\in Z} f(\tilde{w},z)$$

$$\Rightarrow \min_{w\in W} f(w,\tilde{z}) \le \max_{z\in Z} f(\tilde{w},z)$$

$$\Rightarrow \max_{z\in Z}\left(\min_{w\in W} f(w,z)\right) \le \min_{w\in W}\left(\max_{z\in Z} f(w,z)\right)$$

- Lagrange dual problem:
  - Properties
    - * This is a concave problem
    - * The opt solution $d^* \Rightarrow p^*-d^* = \text{gap} \ge 0$

- Slater's Condition

Given that the primal problem is convex. If $f_i(x)<0$ $\forall i=1,\dots,m$, $\exists\, x\in \text{relint}\, D$ the strong duality holds, i.e. gap $=0$.

- Define set

$\mathcal{G}=\{f_1(x),\dots,f_n(x), h_1(x),\dots,h_p(x), f_0(x) \mid x\in D\}$

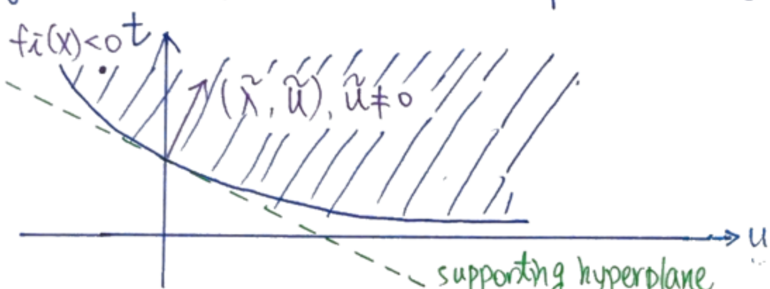
$p^* = \inf\{t \mid (u,v,t)\in \mathcal{G},\ u\le 0, v=0\}$

- Lagrangian

$$L=(\lambda,\nu,1)^T \quad (u,v,t)=\sum_{i=1}^{m}\lambda_i u_i+\sum_{i=1}^{p}\nu_i v_i+t$$

- Dual Problem

$$g(\lambda,\nu)=\inf\left\{(\lambda,\nu,1)^T(u,v,t) \,\middle|\, (u,v,t)\in \mathcal{G}\right\}$$



$$\{(u,t) \mid f_0\le t,\ f_i(x)\le u,\ \exists x\in D\}$$

Since $\tilde{u}\neq 0$, we can have $(\lambda,\nu,1)=\left(\frac{\tilde{\lambda}}{\tilde{u}},\frac{\tilde{\nu}}{\tilde{u}},1\right)$

- Example: Conjugate Function

$$\min f_0(x)$$
$$\text{s.t.}\begin{cases}Ax=b\\ Cx=d\end{cases}$$

- Dual Function

$$\begin{aligned}g(\lambda,\nu)&=\inf_{x\in D}\left[f_0(x)+\lambda^T(Ax-b)+\nu^T(Cx-d)\right]\\&=\inf_{x\in D}\left[\underline{f_0(x)+(A^T\lambda+C^T\nu)^T x}-b^T\lambda-d^T\nu\right]\\&=-f_0^*(y)-b^T\lambda-d^T\nu,\ \text{where } y=-[A^T\lambda+C^T\nu]\end{aligned}$$

- Example: Entropy Maximization

$$\min f_0(x)=\sum_{i=1}^{n}x_i\log x_i$$
$$\text{s.t.}\begin{cases}Ax\le b\\ \mathbf{1}^T x=1\end{cases}$$

- Dual Function

$$\begin{aligned}g(\lambda,\nu)&=-f_0^*(y)-b^T\lambda-d^T\nu,\ \text{where } y=-[A^T\lambda+\nu\cdot\mathbf{1}]\\&=-b^T\lambda-\nu-\sum_{i=1}^{n}e^{-a_i^T\lambda-\nu-1}\end{aligned}$$