Lagrangian Duality. Definitions and classical results

Definition of a Saddle Point: given a function $L(x,u)$, $L: S \times U \rightarrow \mathbb{R}$, defined on two groups of variables: $x \in S \subseteq \mathbb{R}^n$, $u \in U \subseteq \mathbb{R}^n$; then, $(\bar{x},\bar{u}) \in S \times U$ is referred to as a Saddle Point if

$$L(\bar{x},u) \leq L(\bar{x},\bar{u}) \leq L(x,\bar{u})$$

Consider now an optimization problem:

$$(P) \qquad \begin{array}{ll} Min_{\,x} & f(x) \\ F & \left\{ \begin{array}{l} h(x) = 0 \\ g(x) \geq 0 \\ x \in X \end{array} \right. \end{array} \qquad \boxed{\begin{array}{c} \lambda \\ \mu \end{array}} \leftarrow \text{L.M.} \qquad \begin{array}{l} h: \mathbb{R}^n \rightarrow \mathbb{R}^p \\ g: \mathbb{R}^n \rightarrow \mathbb{R}^q \end{array}$$

where no regularity conditions (or constraint qualifications) are assumed a priori and where X may be any type of set:

- For problem (P) the lagrangian function is defined as:

$$\boxed{L(x,\lambda,\mu) \triangleq f(x) - \lambda^\top h(x) - \mu^\top g(x)}$$

- and the dual lagrangian function $w(\lambda,\mu)$ for problem (P) by

$$\boxed{w(\lambda,\mu) \triangleq Inf_{\,x \in X} L(x,\lambda,\mu)}$$

  The domain of definition for w is defined as the points (λ,μ) where w is finite.

$$D \triangleq \{\ (\lambda,\mu) \in \mathbb{R}^p \times \mathbb{R}^q_+ \mid w(\lambda,\mu) > -\infty\ \}$$

- The dual lagrangian problem (LD) for problem (P) is defined as:

$$(LD) \quad \begin{array}{ll} Max & w(\lambda,\mu) \\ & (\lambda,\mu) \in D \end{array}$$

  (remark: the choice of which constraints are dualized may depend on the application or on the convenience of the modeler)

The solution set of the problem defining the dual lagrangian function $w(\cdot,\cdot)$ at a point (λ,μ) will be denoted by $S^*_D(\lambda,\mu)$. The solution set of the dual lagrangian problem (LD) will be denote by S^*_{DL}

Duality and Saddle-Point theorems

Weak Duality theorem

Let x be a feasible point of problem (P) ,$(h(x) = 0,\ g(x) \geq 0,\ x \in X)$. Let (λ,μ) also feasible of (LD), $(\mu \geq 0)$. Then $\boxed{f(x) \geq w(\lambda,\mu)}$.
The difference $g(x,\lambda,\mu) = f(x) - w(\lambda,\mu)$ (≥ 0) is usually called duality gap for problem (P) at (x,λ,μ).

Proof:
$w(\lambda,\mu) = Inf\{\ L(x,\lambda,\mu) \mid x \in X\ \} \ \leq\ f(x) - \lambda^\top h(x) - \mu^\top g(x)\ \leq\ f(x)$
$(\mu \geq 0\ ,\ g(x) \geq 0)$

(Notice that if x^* solves (P) and (λ,μ), $\mu \geq 0$, solves (LD) and there holds that $\mu^\top g(x^*) = 0 \Rightarrow w(\lambda,\mu) = f(x^*)$)

Corollary 1 $f^* = Inf\{\ f(x) \mid x \in F\ \} \geq Sup\{\ w(\lambda, \mu) \mid \mu \geq 0\ \} = w^*$. Duality gap of $(P) \triangleq f^* - w^*$.

Corollary 2 If $f^* = w^*$ then x^* and (λ^*, μ^*) solve (P) and (LD) respectively.

Corollary 3 If $f^* = -\infty \ \Rightarrow\ w(\lambda, \mu) = -\infty\ \ \forall(\lambda, \mu),\ \mu \geq 0$

Corollary 4 If $w^* = \infty \ \Rightarrow\ (P)$ is infeasible $(F = \emptyset)$

The saddle point theorem

A necessary and sufficient condition for problem (P) to have a lagrangian function $L(x, \lambda, \mu)$ with a saddle point $(\bar{x}, \bar{\lambda}, \bar{\mu})$ is:

a) $f^* = f(\bar{x}) = w(\bar{\lambda}, \bar{\mu}) = w^*$ (duality gap=0)

b) $g(\bar{x}) \geq 0\ \ h(\bar{x}) = 0$ $(\bar{x} \in F)$

c) $\bar{\mu}^\top g(\bar{x}) = 0$ (complementarity)

Another Theorem. Saddle points verify Kuhn Tucker conditions

Let $\bar{x} \in F$ a point satisfying K-T for problem (P)

$$\nabla f(\bar{x}) = \left(\frac{\partial g}{\partial x}\right)^\top \bar{\mu} + \left(\frac{\partial h}{\partial x}\right)^\top \bar{\lambda}, \quad \bar{\mu}^\top g(\bar{x}) = 0\ \ ,\ \bar{\mu} \geq 0.$$

Let now: $I(x) = \{\ 1 \leq i \leq g \mid g(x) = 0\ \}$ and assume that problem (P) verifies :

a) f convex.

b) g_i locally convex at $\bar{x}$, $i \in I(\bar{x})$

c) If $\bar{\lambda}_\ell \neq 0 \ \Rightarrow\ h_\ell(\bar{x})$ affine

llavors $(\bar{x}, \bar{\lambda}, \bar{\mu})$ és punt de sella de L per (P).

In particular convex problems have saddle points, which are precisely those who verify K-T conditions

Strong duality theorem (Karlin) (Teo. 6.2.4)

Let X be a convex and non-empty set f , g convex and h affine.

<u>Constraint qualifications.</u>

a) $\exists\tilde{x} \in X$ so that $g(\tilde{x}) > 0$ i $h(\tilde{x}) = 0$

b) $0 \in int\ h(X);\ \ h(X) = \{\ h(x) \mid x \in X\ \}$

If previous conditions are verified then:

1) $f^* = w^*$ (null duality gap)

2) If $f^* < +\infty$ then the maximum of the dual lagrangian function w^* is attained at a point (λ^*, μ^*), $\ w^* = w^*(\lambda^*, \mu^*)$. Moreover, $\mu^* \geq 0$.

3) If $f^* < +\infty$ for x^*, $f^* = f(x^*)$, there holds complementarity, i.e.: $g(x^*)^\top \mu^* = 0$

Previous conditions are <u>sufficient</u> for the existence of a saddle point.
(remark: what happens if there are no equality constraints?)

Duality gaps and solutions of the Dual Lagrangian Problem. Basic relationships

(Minoux-Vajda p.212)

Let $(\bar{\lambda}, \bar{\mu})$ a solution of the Dual Lagrangian Problem.

- If a saddle point exists for problem (P), then there must exist a solution x^* of problem (P) so that $(x^*, \bar{\lambda}, \bar{\mu})$ is a saddle point.
- If a saddle point exists for problem (P) and if x^* is a unique minimum of $L(x, \bar{\lambda}, \bar{\mu})$, then $(x^*, \bar{\lambda}, \bar{\mu})$ is a saddle point and moreover, x^* is an optimal solution of (P).
- If $w(\lambda, \mu)$ is differentiable at $(\bar{\lambda}, \bar{\mu})$ and, additionally, x^* is the unique minimum of $L(x, \bar{\lambda}, \bar{\mu})$, then (P) has a saddle point and $(x^*, \bar{\lambda}, \bar{\mu})$ it is this saddle-point.

(Remark: if there are no saddle points then $w(\cdot, \cdot)$ must be non-differentiable.)

Duality gaps and optimality conditions of the Dual Lagrangian Problem.

Consider the first order optimality conditions of the Dual Lagrangian Problem at an optimal point (λ^*, μ^*):

$$\text{(1stOOC)} \qquad \partial w(\lambda^*, \mu^*) \in \mathsf{N}_{\mathbb{R}^p \times \mathbb{R}^q_+}(\lambda^*, \mu^*)$$

or equivalently:

$$\exists r^* = \begin{pmatrix} r^*_\lambda \\ r^*_\mu \end{pmatrix} \in \partial w(\lambda^*, \mu^*), \text{ so that } r^*_\lambda = 0,\ r^*_\mu \leq 0 \text{ and } {\lambda^*}^\top r^*_\lambda + {\mu^*}^\top r^*_\mu = 0$$

Then, if such an r^* verifying 1st order conditions (1stOOC), is so that:

$$r^* = \begin{pmatrix} r^*_\lambda \\ r^*_\mu \end{pmatrix} = \begin{pmatrix} -h(x^*) \\ -g(x^*) \end{pmatrix}$$

for a solution $x^* \in X$ of the problem defining the dual lagrangian function, i.e.: $x^* \in S^*_D(\lambda^*, \mu^*)$, then problem (P) has a null duality gap.
As a consequence, the following sufficient conditions ensure a null duality gap for problem (P):

$$(\square) \qquad \text{The set } \left\{ \begin{pmatrix} -h(x^*) \\ -g^*(x^*) \end{pmatrix} \ \middle|\ x^* \in S^*_D(\lambda^*, \mu^*) \right\} \text{ is convex}$$

Remark: if, in problem (P), f is pseudoconvex, h and g are affine and X is a closed set the previous condition ($\square$) is satisfied automatically. Also if X is a discret set, then ($\square$) will never hold.