4 Lagrange multipliers and duality for constrained optimization

4.1 Minimization under equality constraints

Let $J : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ be a functional, and $F_i : \Omega \subset \mathbb{R}^n \to \mathbb{R}$, $1 \leq i \leq m < n$ be m functions $C^1(\Omega)$. We consider the problem

$$\inf_{v \in K} J(v)$$

where $K = \{v \in \Omega \ / \ F_i(v) = 0, \ \forall i\}$ is the set of admissible directions.

Definition 4.1 *The constraints are* **regular** *at point $u \in K$ if the vectors $(F_i'(u))_{1 \leq i \leq m}$ are linearly independent. They are* **nonregular** *otherwise.*

Theorem 4.1 *If u is a local minimum of J over K, if the constraints are regular at u, and if J is differentiable at u, then there exist $\lambda_1, \dots, \lambda_m \in \mathbb{R}$, called* **Lagrange multipliers***, such that*

$$J'(u) + \sum_{i=1}^{m} \lambda_i F_i'(u) = 0.$$

The proof is based on the implicit function theorem.

Remark: We can define the **Lagrangian** of the minimization problem of J over K:

$$\mathcal{L}(u, \lambda) = J(u) + \langle \lambda, F(u) \rangle = J(u) + \sum_{i=1}^{m} \lambda_i F_i(u).$$

If the constraints are regular at u, then $\mathcal{L}'_u(u, \lambda) = 0$ and $\mathcal{L}'_\lambda(u, \lambda) = 0$. (u, λ) is then a critical point of the Lagrangian $\mathcal{L}$.

4.2 Saddle points

4.2.1 Definition

Let V and Q be two Hilbert spaces, $U \subset V$ and $P \subset Q$. We consider a Lagrangian $\mathcal{L} : V \times Q \to \mathbb{R}$ (or defined on $U \times P$).

Definition 4.2 *$(u, p) \in U \times P$ is a* **saddle point** *of $\mathcal{L}$ over $U \times P$ if*

$$\mathcal{L}(u, q) \leq \mathcal{L}(u, p) \leq \mathcal{L}(v, p), \quad \forall q \in P, \forall v \in U.$$

Or equivalently,

$$\sup_{q \in P} \mathcal{L}(u, q) = \mathcal{L}(u, p) = \inf_{v \in U} \mathcal{L}(v, p).$$

Theorem 4.2 *If (u, p) is a saddle point of $\mathcal{L} : U \times P \to \mathbb{R}$, then:*

$$\sup_{q \in P} \inf_{v \in U} \mathcal{L}(v, q) = \mathcal{L}(u, p) = \inf_{v \in U} \sup_{q \in P} \mathcal{L}(v, q).$$

Proof
We always have

$$\sup_{q \in P} \inf_{v \in U} \mathcal{L}(v, q) \leq \inf_{v \in U} \sup_{q \in P} \mathcal{L}(v, q).$$

Indeed,

$$\inf_{v\in U} \mathcal{L}(v,\tilde{q}) \le \mathcal{L}(\tilde{v},\tilde{q}) \le \sup_{q\in P} \mathcal{L}(\tilde{v},q).$$

The other inequality is given by the fact that (u,p) is a saddle point:

$$\inf_{v\in U}\sup_{q\in P} \mathcal{L}(v,q) \le \sup_{q\in P} \mathcal{L}(u,q) = \mathcal{L}(u,p) = \inf_{v\in U} \mathcal{L}(v,p) \le \sup_{q\in P}\inf_{v\in U} \mathcal{L}(v,q).$$

$\square$

4.2.2 Inequality constraints

Let $J : V \to \mathbb{R}$ and $F_i : V \to \mathbb{R}$, $1 \le i \le m$, where V is a Hilbert space. Let $K = \{v \in V \ / \ F_i(v) \le 0,\ 1 \le i \le m\}$ be the set of admissible directions. The problem is the following:

$$(\mathcal{P}) \qquad \text{Find } \ u \in K \ \text{ such that } \quad J(u) = \inf_{v\in K} J(v).$$

The Lagrangian corresponding to this minimization problem is:

$$\mathcal{L}(v,q) = J(v) + \sum_{i=1}^{m} q_i F_i(v) = J(v) + \langle q, F(v)\rangle.$$

The admissible set of Lagrange multipliers in the case of inequality constraints is

$$P = \mathbb{R}^m_+ = \{q \in \mathbb{R}^m \ / \ q_i \ge 0,\ \forall 1 \le i \le m\}.$$

Theorem 4.3 *If $(u,p) \in V \times \mathbb{R}^m_+$ is a saddle point of $\mathcal{L}$ on $V \times \mathbb{R}^m_+$, then $u \in K$ and u is a solution of the optimization problem $(\mathcal{P})$.*

Proof
The left inequality is: $\mathcal{L}(u,q) \le \mathcal{L}(u,p),\ \forall q$. Then

$$\sum_{i=1}^{m} (q_i - p_i) F_i(u) \le 0,\ \forall q \in \mathbb{R}^m_+.$$

Then $F_i(u) \le 0$ for all i, and then $u \in K$. $q = 0$ gives $\sum_{i=1}^{m} p_i F_i(u) = 0$.
The right inequality is: $\mathcal{L}(u,p) \le \mathcal{L}(v,p)$, $\forall v \in V$. Then $J(u) \le J(v) + \langle p, F(v)\rangle \le J(v)$, $\forall v \in K$. $\square$

Remark: $\langle p, F(u)\rangle = 0$, and then either $F_i(u) < 0$, and then $p_i = 0$ (inactive constraint $\Rightarrow$ null Lagrange multiplier); or $F_i(u) = 0$, and then $p_i \ge 0$ (active constraint).

Remark: If J and F_i are all differentiable at u, then

$$J'(u) + \sum_{i=1}^{m} p_i F_i'(u) = 0.$$

4.3 Convex optimization: the Kuhn-Tucker theorem

We assume that J and F_i are convex and continuous on V. Then the set of admissible directions K is a convex closed subset of V. Remind that if J is infinite at infinity, then there exists $u \in K$ solution of the minimization problem.

Theorem 4.4 Kuhn-Tucker (necessary and sufficient optimality conditions): *If J and F_i are convex continuous on V, and differentiable on K, if there exists $\bar{v}$ such that $\forall i$, either $F_i(\bar{v}) < 0$, or $F_i(\bar{v}) = 0$ and F_i affine, then u is solution of the minimization problem $(\mathcal{P})$ if and only if $\exists p \in P$ such that (u,p) is a saddle point of $\mathcal{L}$ on $V \times P$, i.e.*

$$F(u) \leq 0, \quad p \geq 0, \quad \langle p, F(u) \rangle = 0, \quad J'(u) + \sum_{i=1}^{m} p_i F_i'(u) = 0.$$

Example: $V = \mathbb{R}^n$, $J(v) = \langle c, v \rangle$ where $c \in \mathbb{R}^n$. Let $F(v) = Av - b$ where A is a $m \times n$ matrix and $b \in \mathbb{R}^m$. We consider the following optimization problem, very well known in linear programming:

$$(\mathcal{P}) \qquad \inf_{Av-b\leq 0} \langle c, v \rangle.$$

We need to assume that $K = \{v \in \mathbb{R}^n \ / \ Av \leq b\}$ is not empty. Then all hypotheses of Kuhn-Tucker theorem are satisfied. And then there exists $p^* \in \mathbb{R}^m$ such that:

$$c + A^T p = 0, \quad \langle p, Au - b \rangle = 0, \quad Au - b \leq 0, \quad p \geq 0.$$

Indeed, $\mathcal{L}(v,p) = \langle c, v \rangle + \langle p, Av - b \rangle = \langle c, v \rangle + \langle A^T p, v \rangle - \langle p, b \rangle$.

Equality constraints: The Kuhn-Tucker theorem can also be applied to **affine equality constraints**, as $F_i(v) = 0 \Leftrightarrow F_i(v) \leq 0$ and $-F_i(v) \leq 0$.

4.4 Duality

Definition 4.3 *Let V and Q be two Hilbert spaces, and $\mathcal{L}$ a Lagrangian defined over a subset $U \times P \subset V \times Q$. We assume that there exists a saddle point $(u,p) \in U \times P$ of $\mathcal{L}$:*

$$\forall q \in P, \quad \mathcal{L}(u,q) \leq \mathcal{L}(u,p) \leq \mathcal{L}(v,p), \quad \forall v \in U.$$

For $v \in U$ and $q \in P$, let us set

$$\mathcal{J}(v) = \sup_{q\in P} \mathcal{L}(v,q), \qquad \mathcal{G}(q) = \inf_{v\in U} \mathcal{L}(v,q).$$

We call the **primal problem** *the minimization problem*

$$\inf_{v\in U} \mathcal{J}(v),$$

and the **dual problem** *the maximization problem*

$$\sup_{q\in P} \mathcal{G}(q).$$

Note that $\mathcal{J}(v)$ can be equal to $+\infty$, and $\mathcal{G}(q)$ can be equal to $-\infty$, but not everywhere as there exists a saddle point.

Theorem 4.5 (duality) *(u,p) is a saddle point of $\mathcal{L}$ over $U \times P$ if and only if*

$$\mathcal{J}(u) = \min_{v\in U} \mathcal{J}(v) = \max_{q\in P} \mathcal{G}(q) = \mathcal{G}(p).$$

Remark: by definition of $\mathcal{J}$ and $\mathcal{G}$, the necessary and sufficient condition for a saddle point is equivalent to:

$$\mathcal{J}(u) = \inf_{v\in U} \left(\sup_{q\in P} \mathcal{L}(v,q) \right) = \sup_{q\in P} \left(\inf_{v\in U} \mathcal{L}(v,q) \right) = \mathcal{G}(p).$$

Proof
Let (u,p) be a saddle point. For $v \in U$, $\mathcal{J}(v) \geq \mathcal{L}(v,p) \geq \mathcal{L}(u,p) = \mathcal{J}(u)$. As $\mathcal{J}(v) \geq \mathcal{J}(u)$ $\forall v$, then $\mathcal{L}(u,p) = \mathcal{J}(u) = \inf_{v \in U} \mathcal{J}(v)$.
Conversely, let us prove that (u,p) is a saddle point. $\mathcal{L}(u,q) \leq \mathcal{J}(u)$, $\forall q \in P$. We also have $\mathcal{L}(v,p) \geq \mathcal{G}(p) = \mathcal{J}(u)$, $\forall v \in U$. Then $\mathcal{L}(u,p) = \mathcal{J}(u)$ and (u,p) is a saddle point. $\square$

Application: We consider the following convex optimization problem: minimize $J(v)$ under the constraints $v \in V$, $F(v) \leq 0$, where J and $F = (F_1, \dots, F_m)$ are convex functions over V. We introduce the Lagrangian

$$\mathcal{L}(v,q) = J(v) + q.F(v), \qquad \forall v \in V, \ \forall q \in \mathbb{R}^m_+.$$

Using the previous notations, for all $v \in V$,

$$\mathcal{J}(v) = \sup_{q \in \mathbb{R}^m_+} \mathcal{L}(v,q) = \begin{cases} J(v) & \text{if } F(v) \leq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

Then the original optimization problem is equivalent to the primal problem:

$$\inf_{v \in V} \mathcal{J}(v).$$

From the previous theorem, it is equivalent to the dual problem:

$$\sup_{q \in \mathbb{R}^m_+} \mathcal{G}(q),$$

where $\mathcal{G}$ is a concave function, and where the constraints are linear. The dual problem is usually much simpler than the primal one.

Example: We consider the minimization of $J(v) = \dfrac{1}{2}\langle Av, v\rangle - \langle b, v\rangle$ under the constraints $v \in \mathbb{R}^n$, and $F(v) = Bv - c \leq 0$, where A is a $n \times n$ symmetric positive definite matrix, $b \in \mathbb{R}^n$, B is a $m \times n$ matrix, and $c \in \mathbb{R}^m$. The Lagrangian is then given by:

$$\mathcal{L}(v,q) = \frac{1}{2}\langle Av, v\rangle - \langle b, v\rangle + \langle q, Bv - c\rangle, \qquad \forall v \in \mathbb{R}^n, \ \forall q \in \mathbb{R}^m_+.$$

The primal minimization problem is then equivalent to the dual maximization problem:

$$\sup_{q \geq 0} \left(-\frac{1}{2}\langle q, BA^{-1}B^T q\rangle + \langle BA^{-1}b - c, q\rangle - \frac{1}{2}\langle A^{-1}b, b\rangle \right).$$

Note that the function to be maximized is still quadratic (concave), but the constraints are much simpler than in the primal problem.