B. Optimisation with constraints

The optimization of a mathematical program with constraints can be described as the optimization of a function (objective) subject to a set of constraints. The function and the constraints can be linear or nonlinear. The constraints can have the same sign or different signs.

Let (P) the program :

Minimize $f(x)$
Subject to :
$g_i(x) = 0$, for $1 \leq i \leq m$,
$h_j(x) \geq 0$, for $1 \leq j \leq p$,
and $x \in IR^n$.

The functions f, gi and hj are supposed to be differentiable.

1. Kuhn and Tucker conditions

Kuhn and Tucker conditions are the generalization of the Lagrange conditions to programs with different constraints signs. These conditions are necessary for each local minimum of (P) on IR^n.

Theorem. Let x* be a local minimum for the mathematical program (P). There exists m real numbers $\lambda_1, \lambda_2, \dots, \lambda_m$ and p positive real numbers $\mu_1, \mu_2, \dots, \mu_p$ such that :

1. $\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla h_j(x^*)$
2. *If* $h_j(x^*) \neq 0$, *then* $\mu_j = 0$.

As Lagrange conditions, Kuhn and Tucker conditions are generally not sufficient. Kuhn and Tucker conditions are sufficient if simultaneously : f is convex, the functions g_i are linear and the functions h_j are concave.

Many optimization methods exist for linear and nonlinear mathematic programs with constraints. Penalties, Barriers and Descent methods are the most used.

2. Penalties and Barriers methods

These methods share the same idea. The function to be optimized contains a ponderation of the original program constraints.
For the mathematical program (P) let the penalties function be:

$\alpha(x) = \sum_{i=1}^{m} |g_i(x)|^q + \sum_{j=1}^{p} |h_j(x^*)|^q$, where q is a positive integer.

Penalties method can be described as follows:

Initialization :

Choose a real number $\varepsilon > 0$ and a starting point x^1.

Choose a penalty parameter $\mu_1 > 0$ and a real number $\beta > 1$.

Let $k = 1$ and go to main step.

Main Step :

1. Starting with x^k, let x^{k+1} be the optimal solution of $f(x) + \mu_k\,\alpha(x)$, go to step 2.
2. If $\mu_k\,\alpha(x^{k+1}) < \varepsilon$, Stop. Else, Let $\mu_{k+1} = \beta\,\mu_k$, replace k by $k+1$ and go to step 1.

The same way we define the barriers function:

$B(x) = \sum_{i=1}^{m} 1/|g_i(x)| + \sum_{j=1}^{p} 1/|h_j(x)|$

Barriers method can be described as follows:

Initialization :

Choose a real number $\varepsilon > 0$ and a starting point x^1.

Choose a barrier parameter $\mu_1 > 0$ and a real number $\beta < 1$ strictly positive.

Let $k = 1$ and go to main step.

Main Step :

1. Starting with x^k, let x^{k+1} be the optimal solution of $f(x) + \mu_k\,B(x)$, go to step 2.
2. If $\mu_k\,B(x^{k+1}) < \varepsilon$, Stop. Else, Let $\mu_{k+1} = \beta\,\mu_k$, replace k by $k+1$ and go to step 1.

One can observe that these two methods are very similar.

Exercise :

Minimize $x_1^2 + x_2^2$

Subject to :

$x_1 + x_2 - 1 = 0$

3. Zoutendijk Method (Descent Method)

The idea is to find a feasible descent direction d starting from a feasible starting point.

Let (P) be the following program :

$$\begin{aligned}&\text{Minimize } f(x)\\&\textit{Subject to :}\\&g_i(x) \le 0, \text{ for } 1 \le i \le m,\\&\text{and } x \in IR^n.\end{aligned}$$

Note that we can transform any equality constraint to a pair of inequalities.

Zoutendijk Method in case of nonlinear constraints can be described :

Initialization :
Choose a staring point x^1 such that $g_i(x) \le 0$, for $1 \le i \le m$.
Let $k = 1$ and **go to main step**.
Main Step :

1. Let $I = \{i : g_i(x_i) = 0\}$ and solve the following program:

   $$\begin{aligned}&\textit{Minimize} && z\\&\textit{Subject to} && \nabla f(x^k)^t d - z \le 0\\& && \nabla g_i(x^k)^t d - z \le 0, \textit{ for } i \in I\\& && -1 \le d_j \le 1, \textit{ for } j = 1,\dots, n\end{aligned}$$

   Let (z_k, d_k) be the optimal solution found. If $z_k = 0$, Stop. The point x_k satisfies the Karush-Kuhn-Tucker conditions. If $z_k < 0$, **go to step 2**.
2. Let λ_k be the optimal solution to the unidimensional program:

   $$\begin{aligned}&\textit{Minimize} && f(x^k + \lambda d_k)\\&\textit{Subject to} && 0 \le \lambda \le \lambda_{max}\end{aligned}$$

   with $\lambda_{max} = \sup\ \{ \lambda : g_i(x^k + \lambda d_k) \le 0, \text{ for } 1 \le i \le m \}$. *Let* $x^{k+1} = x^k + \lambda d_k$, *replace* k *by* $k+1$ *and* ***return to step 1***.

Reference :

NonLinear Programming, Theory and Algorithms, Bazaraa M.S., Sherali H.D and Shetty C.M., second edition, WILEY.