Solution of Multivariable Optimization with Inequality Constraints by Lagrange Multipliers

Consider this problem:

Minimize $f(x)$ where, $x=[x_1\ x_2\ ....\ x_n]^T$

subject to, $g_j(x) \le 0 \qquad j = 1,2,\cdots m$

The *g* functions are labeled *inequality constraints*. They mean that only acceptable solutions are those satisfying these constraints.

Another way to think about an optimization problem with inequality constraint is we are trying to find a solution *within a space bounded by these constraints*.

To start, we need to make distinction between two possibilities for a minimum:

- Interior: *No inequality constraint is active.*
  In this case, a minimum is associated with, $\nabla f(x^*) = 0$
- Exterior: *One or more inequality constraint is active.*
  One possible way to think about this problem is $\nabla f(x^*) \ne 0$ but this is point is the *feasible minimum*.

We can find a solution to the problem by adding non-negative *slack (slackness) variables*, y_j^2 such that,

$$g_j(x) + y_j^{\,2} = 0 \qquad j = 1,2,\cdots m$$

Slack variables are not known beforehand.

The problem now is transformed into:

Minimize $f(x)$ where, $x=[x_1\ x_2\ \ldots.\ x_n]^T$

subject to, $g_j(x) + y_j^{\,2} = 0 \qquad j = 1,2,\cdots m$

In this form, the *Lagrange multiplier* method can be used to solve the above problem by creating this function,

$$Minimize, \quad L = f(x) + \sum_{j=1}^{m} \lambda_j \left(g_j(x) + y_j^{\,2} \right)$$

where, λ_j is the *Lagrange multiplier*. This problem can be solved (*necessary conditions*).

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \quad i = 1,2,\ldots,n$$

$$\frac{\partial L}{\partial \lambda_j} = \left(g_j(x) + y_j^{\,2} \right) = 0 \quad j = 1,2,\ldots,m$$

$$\frac{\partial L}{\partial y_j} = 2\lambda_j y_j = 0 \quad j = 1,2,\ldots,m$$

The total number of equations is *n+2m*, which can be solved simultaneously to obtain the optimal point. The solution will indicate which constraint is active, if any, are associated with the solution.

It may be useful to understand the solution a little better.

- The first set of equations state that the gradient is still zero for the case of an exterior minimum. The gradient now combines the original function and the active constraints.
- The second set of equations ensure that $g_j(x) \le 0$
- The third set of equations indicate either y_j or λ_j is zero.
  - If $\lambda_j=0$, it means that this *constraint is inactive*.
  - If $y_j=0$, it means that this *constraint is active*. ($g_i=0$)

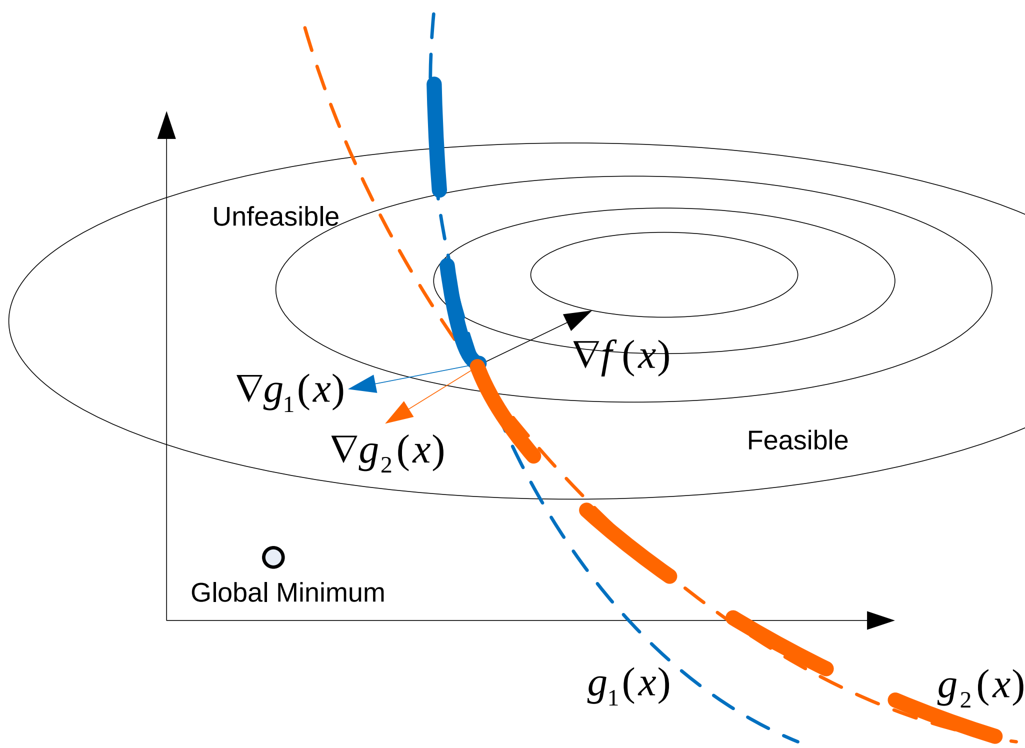
Typically, consider the case when p constraints are active, which means that m-p are inactive. The first equation becomes,

$$-\frac{\partial f}{\partial x_i} = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} \quad i = 1,2,\ldots,p$$

Or,

$$-\nabla f = \sum_{j=1}^{m} \lambda_j \nabla g_j \quad i = 1,2,\ldots,p$$

The figure below may help understand constrained optimization. In this case, the global minimum is outside feasible range. Remember that at minimum slope is zero.

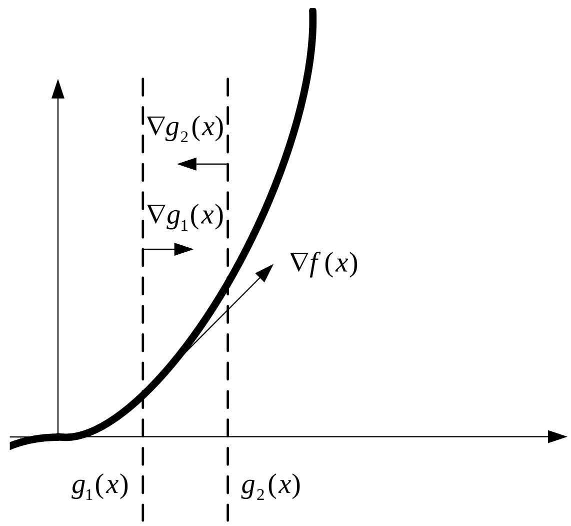


Example 2.7:

Minimize $f(x) = x^3$
Subject to,
$x - 1 \geq 0$
$2 - x \geq 0$

Place the problem in the standard (canonical) form:

Minimize $f(x) = x^3$
Subject to,
$1 - x \leq 0$
$x - 2 \leq 0$



Prepare the solution:
$\nabla f(x) = 3x^2$
$\nabla g_1(x) = -1$
$\nabla g_2(x) = 1$

Condition for minimum:
$3x^2 + \lambda_1(-1) + \lambda_2(1) = 0$
$(1 - x) + {y_1}^2 = 0$
$(x - 2) + {y_2}^2 = 0$
$2\lambda_1 y_1 = 0$
$2\lambda_2 y_2 = 0$

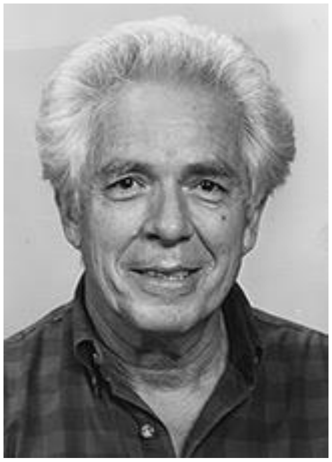
Here we have to explore several possibilities:

• $\lambda_1=0$	$\lambda_2=0$	x=0, which is outside the feasible domain.
• $\lambda_1\neq 0$	$\lambda_2=0$	$y_1=0$ Equation (2) will result in x=1. The first equation means that $\lambda_1=3$ The function is 1
• $\lambda_1=0$	$\lambda_2\neq 0$	$y_2=0$ Equation (3) will result in x=2. The first equation means that $\lambda_2=-12$ The function is 8
• $\lambda_1\neq 0$	$\lambda_2\neq 0$	This case is trivial as both constraints cannot be active in the same time

Homework: 2.61 using the Lagrange Multipliers method. Plot the contour plots and the constraints. Relate your solution to this plot.

Kuhn-Tucker Optimality Conditions

Kuhn and Tucker extended the Lagrange's theory to include classical nonlinear programming problems.



Harold Kuhn (1925-2014), Wikipedia

Albert William Tucker (1905-1995), Wikipedia

Kuhn and Tucker focused on identifying the conditions that when satisfied are related to constrained minimum or,

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \quad i = 1,2,\dots,n$$

$\lambda_j > 0$ $\quad\quad$ $j \in J_1$

The above equations are labeled *Kuhn-Tucker conditions*.

These conditions are necessary but not *necessary* to ensure optimality. They are not however not *sufficient.*

If we limit the discussion to *convex programming problems*, the conditions become both necessary and sufficient.

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \quad i = 1,2,\dots,n$$

$$\lambda_j g_j = 0 \quad j = 1,2,\dots,m$$

$$g_j \leq 0 \quad j = 1,2,\dots,m$$

$$\lambda_j \geq 0 \quad j = 1,2,\dots,m$$

A problem where both the *objective function* and the *constraints* are convex.

Note: The Hessian matrix of a convex function is positive semidefinite.

Constraint Qualifications

We can now that we can solve an optimization problem with equality and inequality constraints as:

Find x, λ, and β vectors such that,

$$\nabla f(x)+\sum_{j=1}^{m}\lambda_j \nabla g_j(x)-\sum_{k=1}^{p}\beta_k \nabla h_k(x)=0$$

$$\begin{array}{lll} g_j(x)\le 0 & j=1,2,\cdots m & \textit{feasibility} \\ h_k(x)=0 & k=1,2,\cdots p & \textit{feasibility} \\ \lambda_j g_j(x)=0 & j=1,2,\cdots m & \\ \lambda_j \ge 0 & j=1,2,\cdots m & \end{array}$$

To be more specific, we need to state that, $\nabla f, \nabla g, and\ \nabla h$ should be *linearly independent*.

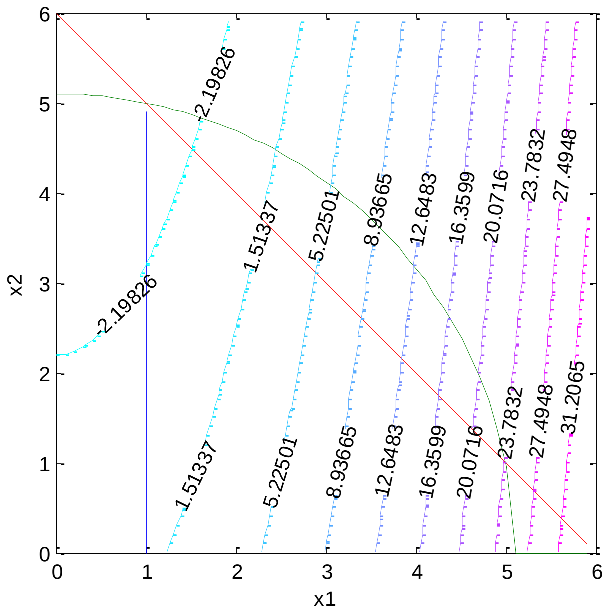
Example 2.8:

Minimize $f(x) = x_1^{\ 2} - x_2$

Subject to, $g_1(x) = -x_1 + 1 \le 0$

$g_2(x) = -26 + x_1^{\ 2} + x_2^{\ 2} \le 0$

$h_1(x) = x_1 + x_2 - 6 = 0$



Graphical inspection shows the minimum is at (1,5).

However, we will solve as if we do not know this.

Prepare the solution,

$$\nabla f(x) = \begin{Bmatrix} 2x_1 \\ -1 \end{Bmatrix}$$

$$\nabla g_1(x) = \begin{Bmatrix} -1 \\ 0 \end{Bmatrix}$$

$$\nabla g_2(x) = \begin{Bmatrix} 2x_1 \\ 2x_2 \end{Bmatrix}$$

$$\nabla h_1(x) = \begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$$

Evaluate the function and constraints

$H(f) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ Positive semidefinite

$H(g_1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ Linear: positive semidefinite

$H(g_2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ Positive definite

Conclusion: We can apply the Kuhn-Tucker conditions:

$$2x_1 + \lambda_1(-1) + \lambda_2(2x_1) - \beta_1(1) = 0$$
$$-1 + \lambda_1(0) + \lambda_2(2x_2) - \beta_1(1) = 0$$
$$-x_1 + 1 \le 0$$
$$-26 + x_1^{\ 2} + x_2^{\ 2} \le 0$$
$$x_1 + x_2 - 6 = 0$$
$$\lambda_1(-x_1 + 1) = 0$$
$$\lambda_2\left(-26 + x_1^{\ 2} + x_2^{\ 2}\right) = 0$$
$$\lambda_j \ge 0 \qquad j = 1,2$$

Here we have to explore several possibilities:

• $\lambda_1=0$ $\lambda_2=0$	Equation (1): $2x_1=\beta_1$ Equation (2): $-1=\beta_1$ Solving these two equations together, $x_1=-0.5$ Equation (5): $x_2=6.5$ (-0.5, 6.5) violates Equation (4), STOP
• $\lambda_1\neq0$ $\lambda_2=0$	Equation (6): $x_1=1$ Equation (5): $x_2=+5$ or -5 (1,-5) violates Equation (5), STOP (1, 5) does not violate any constraint. f=-4
• $\lambda_1=0$ $\lambda_2\neq0$	Equation (7): $x_1^2+x_2^2=26$ Equation (5): $x_1+x_2=6$ Solutions: (1,5) or (5,1) Both solutions do not violate constraints. f=-4 or 24 Choose (1, 5) since we are looking for a minimum.
• $\lambda_1\neq0$ $\lambda_2\neq0$	Equation (6): $x_1=1$ Equation (7): $x_2=+5$ or -5 (1,-5) violates Equation (5), STOP (1, 5) does not violate any constraint. f=-4

Note:

The same solution came from three out of the four cases since the two inequality constraints and the equality constraint intersect at the same point (1, 5).

Homework: 2.64, 2.69, 2.73