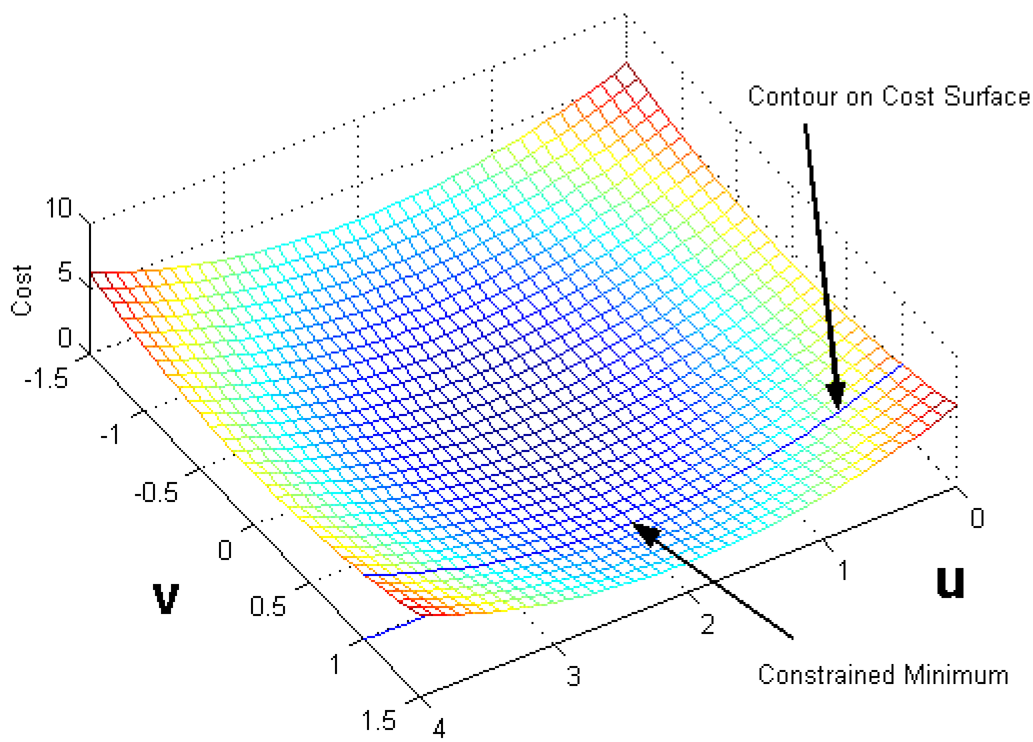
Notes on Applied Math – Constrained Optimization via Lagrange Multipliers
Dr. Fred DePiero

Lagrange multipliers are a powerful and versatile technique for solving cost minimization problems where the permissible solutions are restricted by some secondary relationship (the constraint). There are a wide variety of types of constraints supported by the general methodology of lagrange multipliers. For the purpose of an introduction, only a single, linear, constraint will be considered. Also, the cost function to be minimized will be 2nd order, with two parameters.

The cost function to be minimized is denoted G(u,v) and the constraint has the general form F(u,v) = 0. For example:

$$G(u,v) = (u-2)^2 + v^2 \quad \text{and} \quad F(u,v) = v - 1 = 0$$

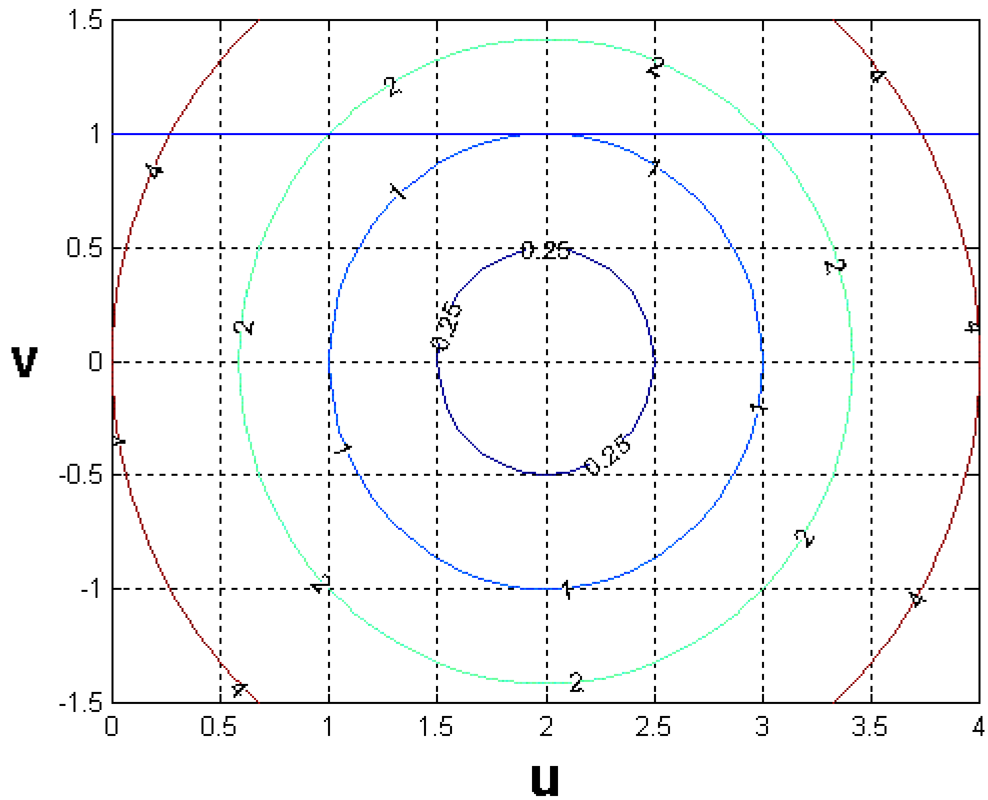


A 2nd order cost function, G(u,v), and linear constraint, F(u,v)=0. The contour of the constraint is shown on the surface of the cost function, and the location of the constrained minimum is indicated.

The goal of the optimization problem is to find an optimal set of values (u_0, v_0) where the cost function, $G(u_0, v_0)$, is minimized; and that obeys the constraint $F(u_0,v_0) = 0$.

Level curves are defined as all points (u,v) where G(u,v) = constant. These play an important role in the constrained optimization. In this example the level curves are circles, centered at (u,v) = (2,0). Because the cost function possesses a global *minimum* (2nd derivatives are positive – it is 'bowl-up'), level curves further from the global minimum correspond to larger values of the cost function.

Referring to the figure below, it can be appreciated that the constrained minimum occurs where a level curve is tangent to the constraint. See $(u_0,v_0)=(2,1)$ below. This tangential point (u_0,v_0) is a location common to both the constraint and to the smallest diameter (lowest cost) level curve. In this case (u_0,v_0) is a point on the level curve $G(u,v)=1$.



Level curves of G(u,v) appear as circles in the example shown. The constraint F(u,v)=0 is a horizontal line. (Level curves may be established for any real value, not just the 0.25, 1, 2 and 4 shown here).

The constrained minimum is computed via the use of gradients. The gradient of $G(u,v)$, $\nabla G(u_0,v_0)$, is a vector field in the u-v plane that points in the direction of the most rapid increase of $G(u,v)$. These vectors are perpendicular to the level curves, because $\nabla G(u,v)$ will not have any component parallel to a level curve.

The gradient of the constraint function, $\nabla F(u_0, v_0)$, is perpendicular to the line of the constraint. This can be appreciated by considering a point (u_1,v_1) where $F(u_1,v_1)=0$. Imagine moving from (u_1,v_1) to a nearby point (u_2,v_2), which is also on the constraint line. In this case the function value $F(u,v)$ remains unchanged. Hence the gradient has no component along the constraint line. Rather the gradient is perpendicular to the constraint line.

Because $\nabla G(u,v)$ is tangent to the level curves, and $\nabla F(u,v)$ is tangent to the constraint, the permissible minimum occurs at (u_0, v_0) where: a) $\nabla G(u_0,v_0)$ and $\nabla F(u_0,v_0)$ align and b) the constraint $F(u_0,v_0)=0$ is satisfied. This alignment condition occurs when

$$\nabla G(u, v) + \lambda \nabla F(u, v) = 0$$

where λ is referred to as a 'Lagrange multiplier'. It serves the purpose of scaling the length of one of the gradients (and reversing direction, as needed) to satisfy the above equation. As such, λ is an additional parameter that must be determined. The constraint $F(u,v) = 0$ provides an additional equation to help solve for all the needed parameters.

To solve for the optimum (u_0, v_0) in the above example, we begin with

$$G(u,v) = u^2 + v^2 - 4u + 4 \quad \text{and} \quad F(u,v) = v - 1 = 0$$

These have gradients

$$\nabla G(u,v) = \begin{bmatrix} \partial G / \partial u \\ \partial G / \partial v \end{bmatrix} = \begin{bmatrix} 2u - 4 \\ 2v \end{bmatrix} \quad \text{and} \quad \nabla F(u,v) = \begin{bmatrix} \partial F / \partial u \\ \partial F / \partial v \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Hence the equations

$$\nabla G(u, v) + \lambda \nabla F(u, v) = 0$$

$$F(u,v) = 0$$

Result in

$$\begin{bmatrix} 2u - 4 \\ 2v \end{bmatrix} + \lambda \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

$$v = 1$$

Which are readily solved to yield

$$(u_0, v_0) = (2, 1) \quad \text{and} \quad \lambda = -2$$