NON LINEAR PROGRAMMING
Prof. Stephen Graves

Consider the example we used to introduce Lagrange multipliers:

$$\text{MIN} \quad f(Q_1, Q_2, Q_3) = \sum_{i=1}^{3} \frac{S_i D_i}{Q_i} + \frac{H_i Q_i}{2}$$

$$\text{s.t.} \quad g(Q_1, Q_2, Q_3) = \sum_{i=1}^{3} \frac{T_i D_i}{Q_i} = K$$

Some definitions:

Gradient of **f**:

$$\nabla f = \left(\frac{\partial f}{\partial Q_1}, \frac{\partial f}{\partial Q_2}, \frac{\partial f}{\partial Q_3} \right)$$

$$= \left(-\frac{S_1 D_1}{Q_1^2} + \frac{H_1}{2}, \ -\frac{S_2 D_2}{Q_2^2} + \frac{H_2}{2}, \ -\frac{S_3 D_3}{Q_3^2} + \frac{H_3}{2} \right)$$

Gradient of **g**:

$$\nabla g = \left(\frac{\partial g}{\partial Q_1}, \frac{\partial g}{\partial Q_2}, \frac{\partial g}{\partial Q_3} \right)$$

$$= \left(-\frac{T_1 D_1}{Q_1^2}, \ -\frac{T_2 D_2}{Q_2^2}, \ -\frac{T_3 D_3}{Q_3^2} \right)$$

Directional derivatives:

Let $\underline{\mathbf{x}} = (x_1, x_2, x_3)$ denote a direction. The directional derivative for **f** and **g** are given by:

$$\frac{d\,f}{d\,\underline{x}} = \sum_{i=1}^{3} \left(\frac{\partial f}{\partial Q_i}\right) x_i = \sum_{i=1}^{3} \left(-\frac{S_i D_i}{Q_i^2} + \frac{H_i}{2}\right) x_i$$

$$\frac{d\,g}{d\,\underline{x}} = \sum_{i=1}^{3} \left(\frac{\partial g}{\partial Q_i}\right) x_i = \sum_{i=1}^{3} \left(-\frac{T_i D_i}{Q_i^2}\right) x_i$$

Note: the directional derivative is the "dot product" of the gradient and the direction vector.

For a given point, say (Q_1, Q_2, Q_3), <u>what is the direction of steepest ascent</u> for the objective function **f**? That is, what direction provides the largest value for the directional derivative? The direction of steepest ascent will be the solution to the following optimization problem:

$$\text{MAX} \quad \frac{d\,f}{d\,\underline{x}} = \sum_{i=1}^{3} \left(-\frac{S_i D_i}{Q_i^2} + \frac{H_i}{2}\right) x_i$$

$$\text{s. t.} \qquad x_1^2 + x_2^2 + x_3^2 = 1$$

By using a Lagrange multiplier to solve this problem, you can show that the direction of steepest ascent is given by

$$\underline{\mathbf{x}}^* = (x_1, x_2, x_3) = \nabla f / \|\nabla f\|$$

That is, the "best" direction is the (normalized) gradient; for our example, the direction of steepest ascent (not normalized) is

$$\underline{\mathbf{x}}^* = \left(-\frac{S_1 D_1}{Q_1^2} + \frac{H_1}{2}\,,\; -\frac{S_2 D_2}{Q_2^2} + \frac{H_2}{2}\,,\; -\frac{S_3 D_3}{Q_3^2} + \frac{H_3}{2}\right)$$

Note that since the actual problem is a minimization, we actually want the <u>direction of steepest descent</u>, which would just be the negative of the gradient.

Now suppose that the given point (Q_1, Q_2, Q_3) is feasible; that is, it satisfies the constraint: $g(Q_1, Q_2, Q_3) = K$

What is the best feasible direction?

The best feasible direction (for ascent) will be the solution to the following optimization problem:

$$\text{MAX} \quad \frac{d\,f}{d\,\underline{x}} = \sum_{i=1}^{3} \left(-\frac{S_i D_i}{Q_i^2} + \frac{H_i}{2} \right) x_i$$

$$\text{s. t..} \quad x_1^2 + x_2^2 + x_3^2 = 1$$

$$\frac{d\,g}{d\,\underline{x}} = \sum_{i=1}^{3} \left(-\frac{T_i D_i}{Q_i^2} \right) x_i = 0$$

By using two Lagrange multipliers to solve the problem, you can find that the best feasible direction (called the reduced gradient) is given by:

$$\underline{x}^* = \left(x_1^*, x_2^*, x_3^* \right) = \nabla f - \left(\frac{\nabla f \bullet \nabla g}{\nabla g \bullet \nabla g} \right) \nabla g$$

where $\nabla f \bullet \nabla g = \sum_{i=1}^{3} \left(\frac{\partial f}{\partial Q_i} \frac{\partial g}{\partial Q_i} \right)$, and $\nabla g \bullet \nabla g = \sum_{i=1}^{3} \left(\frac{\partial g}{\partial Q_i} \frac{\partial g}{\partial Q_i} \right)$

The reduced gradient can then be used as a search direction to improve the current solution. To see that the reduced gradient is a feasible direction, we note that

$$\underline{x}^* \bullet \nabla g = \nabla f \bullet \nabla g - \left(\frac{\nabla f \bullet \nabla g}{\nabla g \bullet \nabla g} \right) \nabla g \bullet \nabla g = \underline{\mathbf{0}}$$

Besides determining the reduced gradient, an algorithm would also need to determine the "step size:" namely how far to move along the reduced gradient. An algorithm stops when the reduced gradient equals (approximately) the zero vector.