Constrained Minimization with Lagrange Multipliers

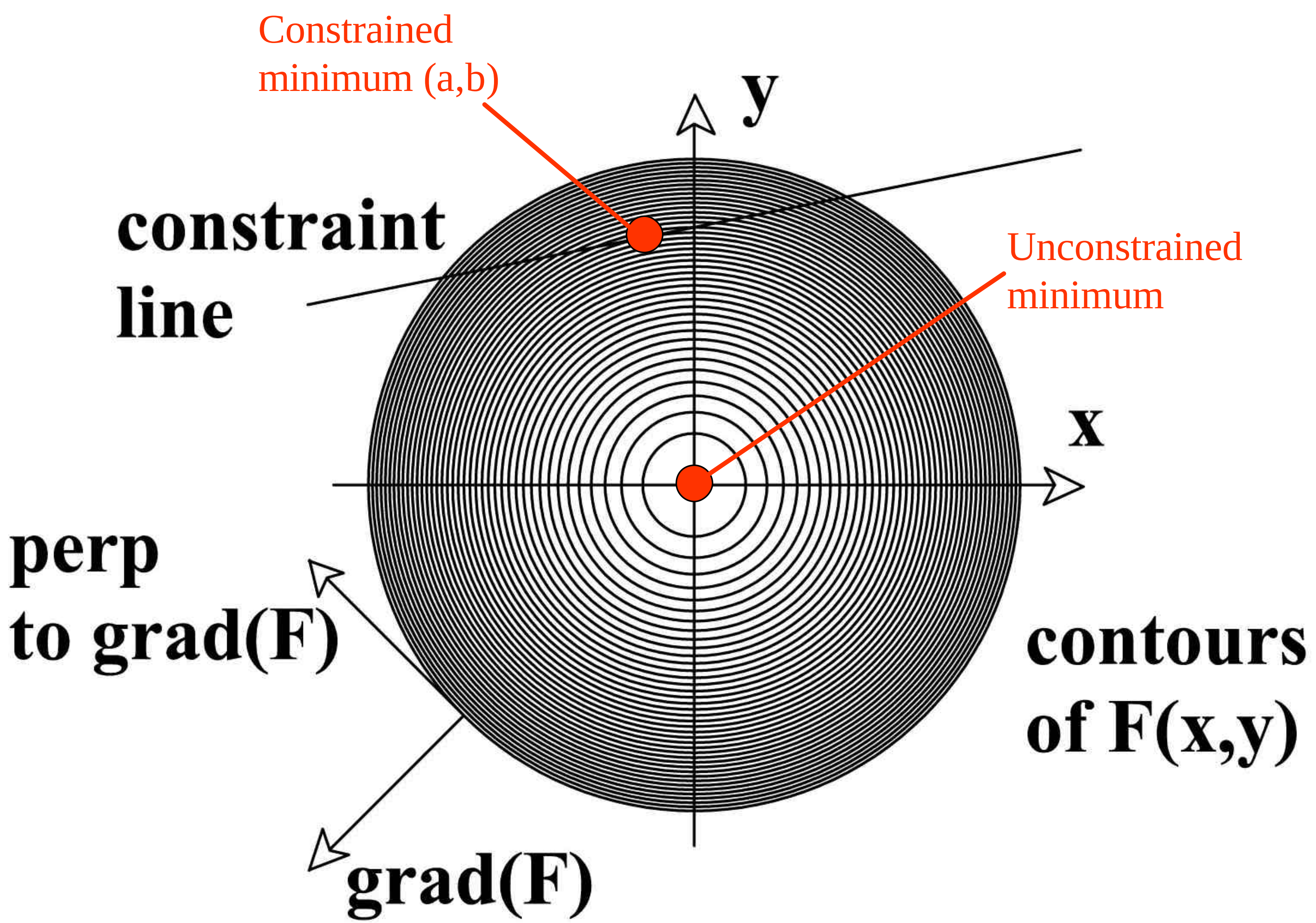
We wish to minimize, i.e. to find a local minimum or stationary point of

$$F(x, y) = x^2 + y^2 \qquad (1)$$

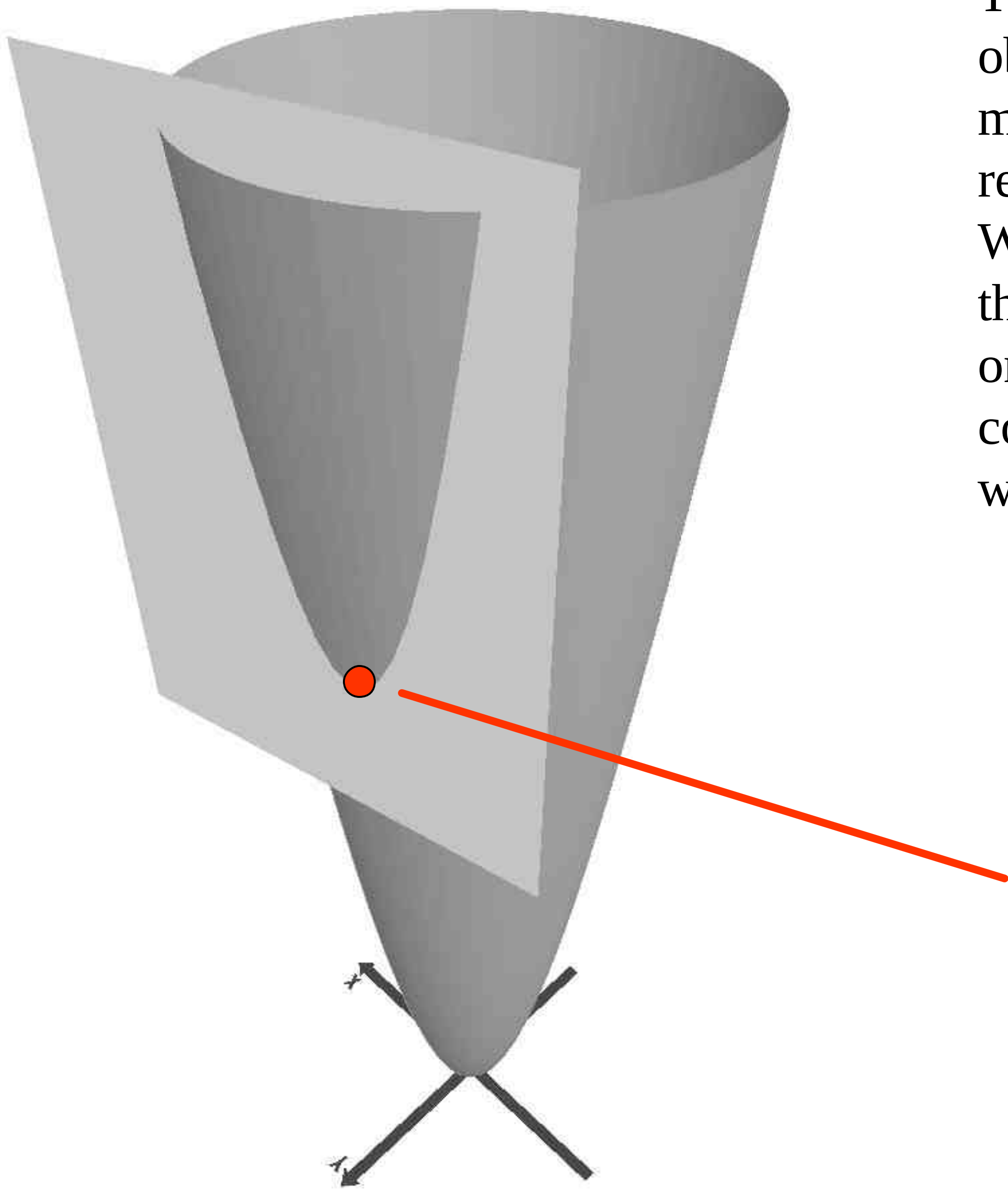
Subject to the equality constraint,

$$\begin{aligned} y &= 0.2x + 5, \quad \text{or} \\ g(x, y) &= y - 0.2x - 5 = 0 \end{aligned} \qquad (2)$$

Look at the surface defined by F(x,y) and sketch the contours,


Constrained
minimum (a,b)
y
constraint
line
Unconstrained
minimum
x
perp
to grad(F)
contours
of F(x,y)
grad(F)

Where the line is the constraint line $g(x,y)=0$ and (a,b) will be the desired minimum point. In 3D, looking from the “north” or $+y$ direction, The bowl function is the objective function, to be minimized. The plane represents the constraint. When we limit or constraint the search for the minimum only to points that satisfy the constraint, by intuition we will get the indicated point.



The minimum of F(x,y) along g(x,y)=0 at (a,b) is just the point where the constraint line is *tangent* to a level curve or contour line. It is also at the point where the constraint line is *perpendicular* to the direction of the gradient, or the direction of maximum slope. The gradient at a point is,

$$grad(F) = \frac{dF}{dx}\mathbf{i} + \frac{dF}{dy}\mathbf{j} \qquad (3)$$

This gives the direction in which the rate of change of the function value is *maximum.* A line in the plane, through (a,b) and parallel to the gradient is,

$$\frac{dF}{dx}(y-b) - \frac{dF}{dy}(x-a) = 0 \qquad (4)$$

A line through (a,b) perpendicular to the gradient, or tangent to the level curve is,

$$\frac{dF}{dx}(x-a) + \frac{dF}{dy}(y-b) = 0 \qquad (5)$$

At the desired minimum point, the line tangent to the level curve must be coincident with the constraint line

The constraint line (2) can be rewritten as,

$$g(x,y) = \frac{dg}{dx}(x-a) + \frac{dg}{dy}(y-b) = 0 \qquad (6)$$

The lines corresponding to equations (5) and (6) must be coincident, that is they must be the *same* lines.

$$\begin{aligned} \frac{dF}{dx}(x-a) + \frac{dF}{dy}(y-b) &= 0 \\ \frac{dg}{dx}(x-a) + \frac{dg}{dy}(y-b) &= 0 \end{aligned} \qquad (7)$$

Therefore, one equation must be a multiple of the other, or the coefficients must be proportional.

$$\begin{aligned} \frac{dF}{dx}(a,b) &= l'\frac{dg}{dx}(a,b) \\ \frac{dF}{dy}(a,b) &= l'\frac{dg}{dy}(a,b) \end{aligned} \qquad (8)$$

Or,

$$\frac{dF}{dx}(a,b) + l\,\frac{dg}{dx}(a,b) = 0$$

$$\frac{dF}{dy}(a,b) + l\,\frac{dg}{dy}(a,b) = 0 \qquad (9)$$

where, $l = -l'$

If we take the two equations in (9), plus the constraint equation (2) together,

$$\frac{dF}{dx}(a,b) + l\,\frac{dg}{dx}(a,b) = 0$$

$$\frac{dF}{dy}(a,b) + l\,\frac{dg}{dy}(a,b) = 0 \qquad (10)$$

$$g(a,b) = 0$$

We now have 3 equations in 3 unknowns, (a, b, l)

Which we can solve. We are generally not interested in the explicit value for lambda, it is necessary but only an intermediate step. So we solve the equations, keeping the (a,b) and throwing away the lambda, the *Lagrange Multiplier*.

For our numerical example,

$$
\begin{aligned}
&2a + l\,(-0.2) = 0 \\
&2b + l\,(1) = 0 \qquad (11) \\
&b - 0.2a - 5 = 0
\end{aligned}
$$

In matrix form,

$$
\begin{bmatrix} 2 & 0 & -0.2 \\ 0 & 2 & 1 \\ -0.2 & 1 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \\ l \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} \qquad (12)
$$

Solving,

$$
\begin{bmatrix} a \\ b \\ l \end{bmatrix} = \begin{bmatrix} -0.9615 \\ 4.8077 \\ -9.6154 \end{bmatrix} \qquad (13)
$$

Note that equations (1) are analogous to what we get in Least Squares (observations only) when we differentiate the *augmented* objective function,

$$\Phi' = v_1^2 + v_2^2 + 2l\,(\text{condition equation}) \Rightarrow \text{minimum} \qquad (14)$$

$$\frac{d\Phi'}{dv_1} = 0$$

$$\frac{d\Phi'}{dv_2} = 0 \qquad (15)$$

$$\frac{d\Phi'}{dl} = 0$$

v_1 and v_2 here are playing the role of the (x,y) or (a,b) unknowns in the derivation. Note that you can solve simultaneously for the v's and lambda's (text uses k). Or you can solve for the v's in terms of the lambda's, then plug into the condition equations (eliminating the v's leaving only the lambda's). Then you solve only for the lambda's, and solve again for the v's from your elimination equations.

Reference: Fulks, Advanced Calculus, p. 264, Lib# 517 F957a Ed. 1