

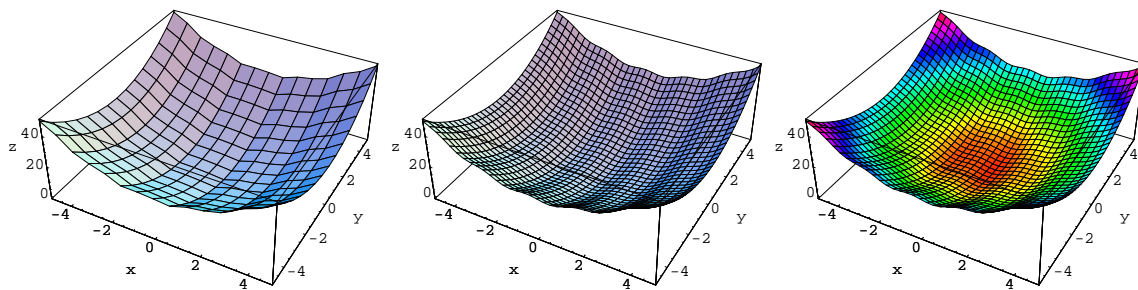
4452 Mathematical Modeling Lecture 3: Unconstrained Multivariable Optimization

Example Say we want to find the maximum of a function in three dimensions.

$$z = f(x, y).$$

First, we want to think of ways to visualize such a function. We can get simple three dimensional plots using *Mathematica*:

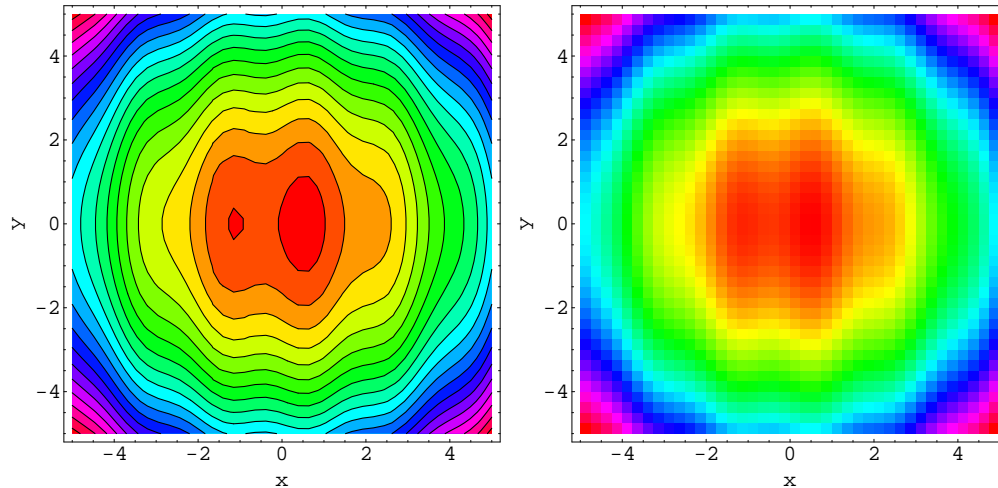
```
f[x_, y_] = x^2 + y^2 - Cos[Pi x]
Plot3D[f[x, y], {x, -5, 5}, {y, -5, 5}, AxesLabel -> {x, y, z}]
Plot3D[f[x, y], {x, -5, 5}, {y, -5, 5}, AxesLabel -> {x, y, z},
  PlotPoints -> 60]
Plot3D[f[x, y], {x, -5, 5}, {y, -5, 5}, AxesLabel -> {x, y, z},
  PlotPoints -> 60, ColorFunction -> Hue]
```



But, we have to be careful that we are seeing all the detail in the graph! It is hard to detect the wave of the cosine in the first graph. A useful way to show location of extrema on a graph is to use the option `ColorFunction -> Hue`.

Other, possibly more useful ways to show a three dimensional function graphically are with level curves or density plots. Level curves (or contour plots) are curves in the xy -plane for which $f(x, y) = k$, where k is a constant. Density plots shade the xy -plane based on the value of $f(x, y)$. In *Mathematica*, these are produced by:

```
ContourPlot[f[x, y], {x, -5, 5}, {y, -5, 5}, Contours -> 20, PlotPoints -> 40,
  FrameLabel -> {x, y}, ColorFunction -> Hue]
DensityPlot[f[x, y], {x, -5, 5}, {y, -5, 5}, PlotPoints -> 40,
  FrameLabel -> {x, y}, ColorFunction -> Hue]
```



These latter plots are potentially more useful in higher dimension. As always, it becomes difficult to visualize as dimension goes up,

$$w = f(x, y, z), w = f(x, y, z, t), \text{ etc.}$$

and, of course, optimization problems are not constricted to optimizing a function of two variables.

For unconstrained optimization, we have no constraints, which essentially means our functions are defined for all \mathbb{R}^n in the space we are trying to optimize over.

Theorem If f has a local max/min at a point $(\bar{x}_1, \dots, \bar{x}_n)$ and the first order partial derivatives of f exist there, then

$$\begin{aligned} f_{x_1}(\bar{x}_1, \dots, \bar{x}_n) &= \frac{\partial f}{\partial x_1}(\bar{x}_1, \dots, \bar{x}_n) = 0, \\ &\vdots \\ f_{x_n}(\bar{x}_1, \dots, \bar{x}_n) &= \frac{\partial f}{\partial x_n}(\bar{x}_1, \dots, \bar{x}_n) = 0. \end{aligned}$$

We also need to know whether we have a max or a min. For function of two variables there is the Second Derivative Test. For functions of more than two variables, tests are difficult to obtain, and we need to rely on the analysis of the physical nature of the problem to determine whether we have found a max or a min (and we could always use this sort of analysis for a function of two variables as well!).

Second Derivative Test. Suppose that the second partial derivatives of f are continuous on a disk with center $(\bar{x}_1, \dots, \bar{x}_n)$, and suppose that

$$\begin{aligned} f_{x_1 x_1}(\bar{x}_1, \bar{x}_2) &= \frac{\partial^2 f}{\partial x_1^2}(\bar{x}_1, \bar{x}_2) > 0, \\ &\vdots \\ f_{x_1 x_1}(\bar{x}_1, \bar{x}_2) &= \frac{\partial^2 f}{\partial x_1^2}(\bar{x}_1, \bar{x}_2) < 0. \end{aligned}$$

Let

$$D = D(\bar{x}_1, \bar{x}_2) = f_{x_1 x_1}(\bar{x}_1, \bar{x}_2) f_{x_2 x_2}(\bar{x}_1, \bar{x}_2) - [f_{x_1 x_2}(\bar{x}_1, \bar{x}_2)]^2$$

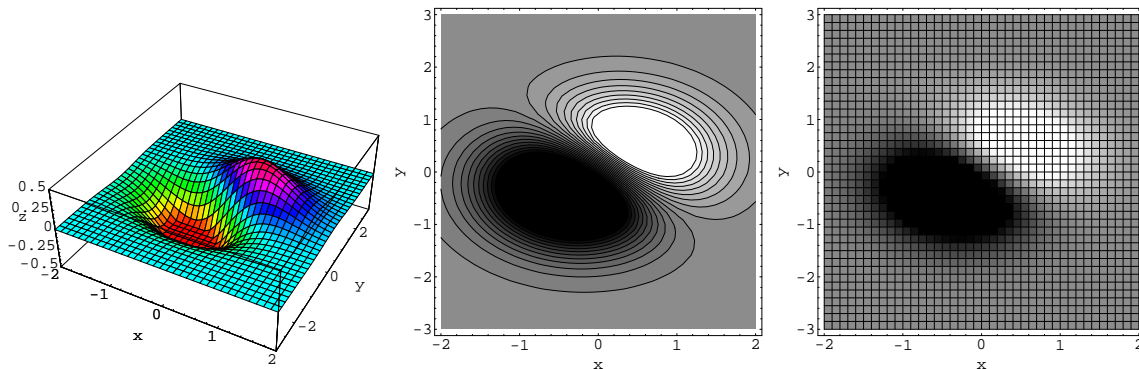
Then if

- (a) $D > 0$ and $f_{x_1 x_1} > 0$, then $f(\bar{x}_1, \bar{x}_2)$ is a local min,
- (b) $D > 0$ and $f_{x_1 x_1} < 0$, then $f(\bar{x}_1, \bar{x}_2)$ is a local max,
- (c) $D < 0$, then $f(\bar{x}_1, \bar{x}_2)$ is a saddle point.

Example Find the maximum and minimum of $f(x, y) = (x - y - 1/4)e^{-x^2 - y^2}$.

First we will get some plots of the situation. We can actually use *Mathematica* to solve this for us easily.

```
f[x_, y_] = (x + y - 1/4) Exp[-x^2 - y^2]
Plot3D[f[x, y], {x, -2, 2}, {y, -3, 3}, AxesLabel -> {x, y, z},
  PlotPoints -> 35, ColorFunction -> Hue, PlotRange -> {-1/2, 1/2}]
ContourPlot[f[x, y], {x, -2, 2}, {y, -3, 3}, Contours -> 20, PlotPoints -> 40,
  FrameLabel -> {x, y}]
DensityPlot[f[x, y], {x, -2, 2}, {y, -3, 3}, PlotPoints -> 40,
  FrameLabel -> {x, y}, Mesh -> True]
Solve[{D[f[x, y], y] == 0, D[f[x, y], x] == 0}, {x, y}]
```



From the three dimensional plot it is easy to identify a maximum and minimum. It looks like the minimum occurs in the third quadrant of the xy -plane, and the maximum occurs in the first quadrant. The contour plot and density plot convey the same information, but it is difficult to tell which is the max and which is the min. You need to know that *Mathematica* shades in such a way that the higher values of $f(x, y)$ are lighter.

Aside: *Mathematica* requires a special package to label the curves in a contour plot. See <http://support.wolfram.com/applicationpacks/unsupported/extend/>.

To solve analytically:

$$\begin{aligned} f(x, y) &= (x - y - 1/4)e^{-x^2 - y^2} \\ f_x(x, y) &= \frac{1}{2}(2 + x - 4x^2 - 4xy)e^{-x^2 - y^2} \end{aligned}$$

$$f_y(x, y) = \frac{1}{2}(2 + y - 4y^2 - 4xy)e^{-x^2 - y^2}$$

$$\text{Solve: } \begin{cases} \frac{1}{2}(2 + x - 4x^2 - 4xy)e^{-x^2 - y^2} = 0 \\ \frac{1}{2}(2 + y - 4y^2 - 4xy)e^{-x^2 - y^2} = 0 \end{cases} \text{ for } (x, y)$$

$$(x, y) = \left(\frac{1}{16} - \frac{\sqrt{65}}{16}, \frac{1}{16} - \frac{\sqrt{65}}{16}\right) = (-0.441, -0.441)$$

$$(x, y) = \left(\frac{1}{16} + \frac{\sqrt{65}}{16}, \frac{1}{16} + \frac{\sqrt{65}}{16}\right) = (0.566, 0.566)$$

Since our plot told us the maximum occurs in the first quadrant, we know we have a max of $f(x, y)$ is $f(0.566, 0.566) = 0.465$. The minimum is $f(-0.441, -0.441) = -0.767$. We could also use the second derivative test to verify these are maximums and minimums.

Sensitivity Analysis

We proceed to analyse the sensitivity of the two dimensional optimization problem as we did before, by leaving the parameter we want to check the sensitivity of unspecified and solving for the optimal values in terms of this parameter.

Example Find the sensitivity of the values of x and y which produce a minimum to the parameter α in the function $f(x, y) = (x - \alpha y - 1/4)e^{-x^2 - y^2}$ near $\alpha = 1$.

To solve analytically:

$$f(x, y) = (x - \alpha y - 1/4)e^{-x^2 - y^2}$$

$$f_x(x, y) = \frac{1}{2}(2 + x - 4x^2 - 4\alpha xy)e^{-x^2 - y^2}$$

$$f_y(x, y) = \frac{1}{2}(2\alpha + y - 4\alpha y^2 - 4xy)e^{-x^2 - y^2}$$

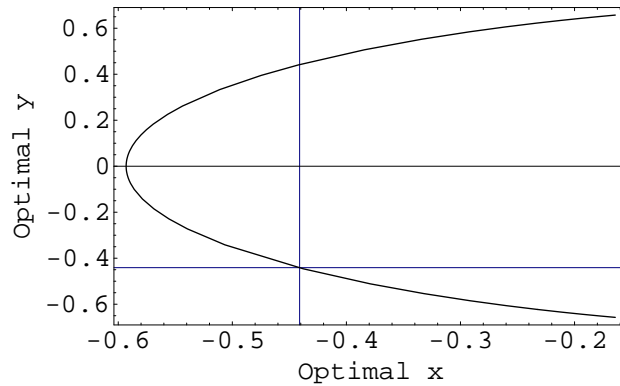
$$\text{Solve: } \begin{cases} \frac{1}{2}(2 + x - 4x^2 - 4\alpha xy)e^{-x^2 - y^2} = 0 \\ \frac{1}{2}(2\alpha + y - 4\alpha y^2 - 4xy)e^{-x^2 - y^2} = 0 \end{cases} \text{ for } (x, y)$$

$$(x, y) = \left(\frac{1}{8(1 + a^2)} \left(1 - \sqrt{33 + 32a^2}\right), \frac{a}{8(1 + a^2)} \left(1 - \sqrt{33 + 32a^2}\right)\right)$$

You get another solution of the form $1 + \sqrt{\quad}$. We can check that the one we wrote is the solution corresponding to the minimum by setting $a = 1$ and recovering the point $(-0.441, -0.441)$.

So we now have the optimal values of x and y (the values which produce a minimum) in terms of the parameter α . In the book, they plot x vs α and y vs α , which is fine, but I look at what we have and see that we have a parametric curve in the xy -plane with parameter α . So I think we should plot this curve instead of the other two.

```
f[x_, y_] = (x + a y - 1/4) Exp[- x^2 - y^2]
sol = Solve[{D[f[x, y], y] == 0, D[f[x, y], x] == 0}, {x, y}]
x = x /. x -> Part[sol, 1, 1, 2]
y = y /. y -> Part[sol, 1, 2, 2]
ParametricPlot[{x, y}, {a, -4, 4}, Frame -> True,
  FrameLabel -> {"Optimal x", "Optimal y"}, GridLines -> {{-0.441}, {-0.441}}]
```

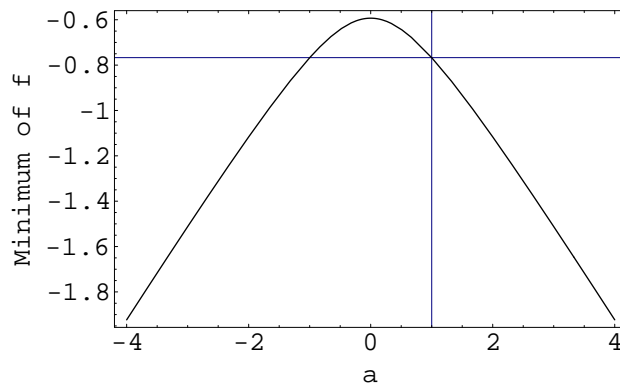


The $\alpha = 1$ point is identified on the curve. We traverse the curve counterclockwise as α increases. Therefore, we see that increasing α from its assumed value of 1 will cause the optimal x to increase, and the optimal y to decrease.

This behaviour would be different if the assumed value of α were -1 . Then, the minimum would be at the point $(-0.441, 0.441)$, and increasing α would cause the optimal value of x to decrease, and the optimal value of y to decrease.

We can plot the optimal value of the function, which in this case is just the minimum value of the function, as a function of α . Again, the $\alpha = 1$ result is identified on the graph.

```
Plot[f[x, y], {a, -4, 4}, Frame -> True, FrameLabel -> {"a", "Minimum of f"},
  GridLines -> {{1}, {-0.767}}, Axes -> False]
```



Increasing α will reduce the value of the minimum.

We want to talk about sensitivities, so let's calculate those quantities.

The sensitivity of the optimal x value with respect to the parameter α is given by:

$$S(x, \alpha) = \frac{dx}{d\alpha} \cdot \frac{\alpha}{x} = \frac{2a^2(-17 - 16a^2 + \sqrt{33 + 32a^2})}{(1 + a^2)\sqrt{33 + 32a^2}(-1 + \sqrt{33 + 32a^2})}$$

The sensitivity of the optimal y value with respect to the parameter α is given by:

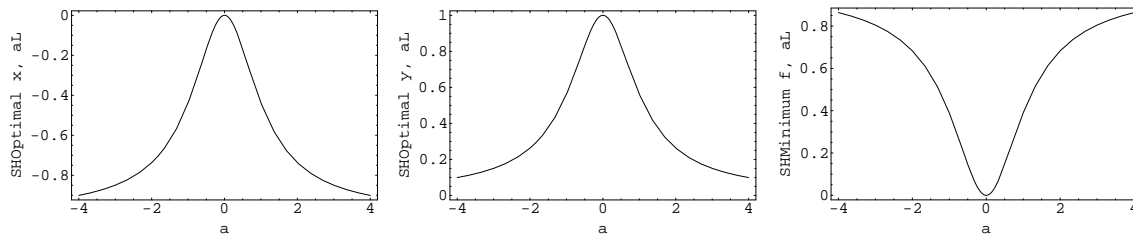
$$S(y, \alpha) = \frac{dy}{d\alpha} \cdot \frac{\alpha}{y} = \frac{33 - \sqrt{33 + 32a^2} + a^2 (31 + \sqrt{33 + 32a^2})}{(1 + a^2) \sqrt{33 + 32a^2} (-1 + \sqrt{33 + 32a^2})}$$

The sensitivity of the optimal value of the function, or the minimum of the function in this case, with respect to the parameter α is given by:

$$S(f, \alpha) = \frac{df}{d\alpha} \cdot \frac{\alpha}{f} = \frac{a^2 (33 + 32a^2 - \sqrt{33 + 32a^2})}{(1 + a^2) \sqrt{33 + 32a^2} (1 + \sqrt{33 + 32a^2})}$$

All of these quantities were easily calculated by *Mathematica* using (and TeXForm for output!):

```
Simplify[D[x, a] a/x]
Simplify[D[y, a] a/y]
Simplify[D[f[x, y], a] a/f[x, y]]
```



The sensitivity of the optimal x to α when $\alpha = 1$ is $S(x, 1) \sim -0.4 = -2/5$, which means that a 5% increase in α will result in a 2% decrease in the optimal x .

The sensitivity of the optimal y to α when $\alpha = 1$ is $S(y, 1) \sim 0.6 = 3/5$, which means that a 5% increase in α will result in a 3% increase in the optimal y .

The sensitivity of the minimum to α when $\alpha = 1$ is $S(f, 1) \sim 0.4 = 2/5$, which means that a 5% increase in α will result in a 2% increase in the value of the minimum of $f(x, y)$.