Night 4: Optimization and Gradient Ascent

Quantitative Engineering Analysis

Spring 2019

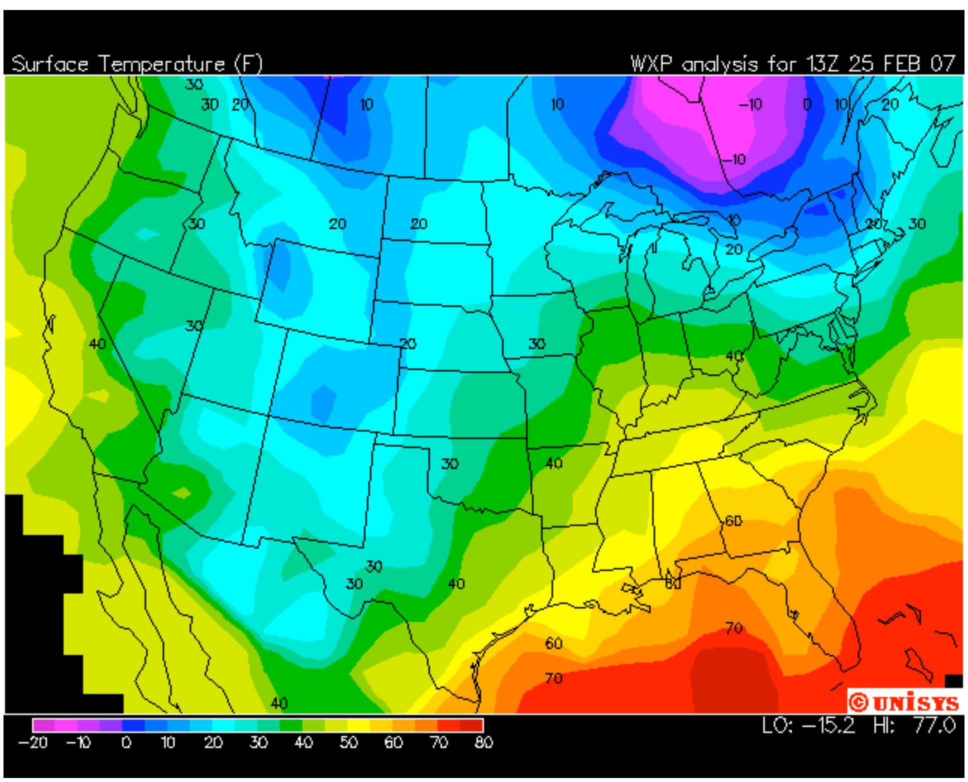
1 *Learning Goals*

By the end of this assignment, you should feel confident with the following:

- Extending the idea of optimization
- Partial derivatives and the gradient
- Unconstrained optimization: concept and the method of steepest ascent

2 *Conceptual Exercise: The Leisure Seeker [1 Hr]*

The map below gives the temperature across the United States on a certain winter day. Regions of the same color have the same temperature: violet represents the coldest areas, and temperatures rise as the colors traverse the spectrum from indigo to blue to green to yellow to orange to red.



Locate Chicago on the map and mark it with a dot. The weather in Chicago is freezing in winter, so a resident of the city decides to embark on a journey in search of the sun. From Chicago, she wants to travel in the direction in which temperature rises most quickly.

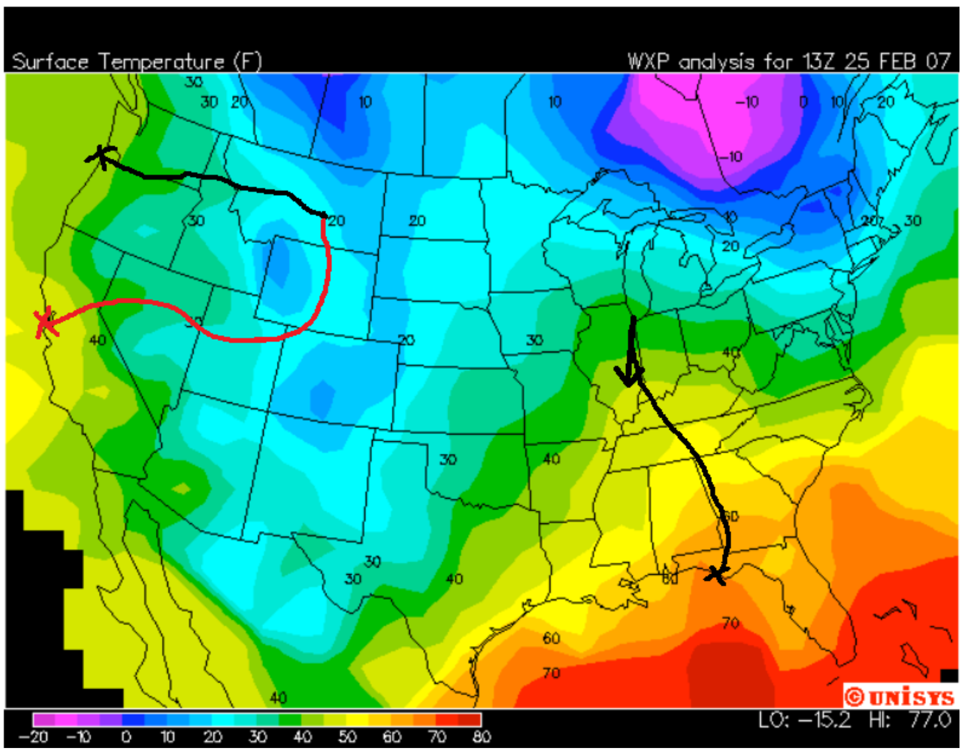
Exercise (1) In which direction should she go? Draw an arrow, starting at Chicago, that indicates this direction.

As her journey proceeds, she decides to keep traveling in the direction in which the weather warms up most quickly: wherever she is at any moment, she moves in the direction of fastest temperature rise.

Exercise (2) Make a rough sketch of the route she takes.

Solution:

See Fig. **??**.



Exercise (3) Where does she end up, assuming that she doesn't leave the United States?

Solution:

Near Panama City, FL

Exercise (4) What would happen if her friend started in Billings, Montana? Where would he end up?

Solution:

Hard to know exactly without more isotherms (contours of constant temperature), but maybe the SF Bay Area or the Portland area.

3 *Readings, Videos, and Conceptual Questions - Partials and Gradients [3 Hrs]*

At this link you will find a set of readings and videos about partial derivatives, the gradient, and the Hessian. Some of this is stuff you already dealt with in the boats and faces modules; some of it is new. Read the text and/or watch these videos, and then write short qualitative answers to the following questions:

Exercise (5) What is meant by f_x? By $\frac{\partial^2 f}{\partial x^2}$? By $D_u f$? By ∇f?

Solution:

The partial derivative with respect to x, meaning the slope in the x direction (where every other independent variable is held constant); the second partial derivative with respect to x, meaning the curvature in the x direction (where every other independent variable is held constant); the directional derivative with respect to vector u, meaning the slope in the u direction; the gradient, meaning the maximum slope in its corresponding direction.

Exercise (6) In Stewart, the idea of gradient is discussed primarily in two dimensional and three dimensional "physical" spaces, using $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and so forth. But more generally, you can have a gradient of a function of any number of variables. Give a real-world example of a gradient for a situation that involves more than 3 variables.

Solution:

Any kind of optimization problem could have a gradient of the objective function in many dimensions (decision variables).

Exercise (7) "The gradient always points in the direction of fastest increase." Can you think of physical examples where there is *more than one* direction of fastest increase? What's going on here?

Solution:

At the bottom of a cone, the increase is equally fast in all directions.

Exercise (8) "The gradient is always normal to level curves/surfaces." Explain.

Solution:

A differential surface area at any given point on a surface is flat, an can be defined by the line where it intersects with the level curve and a perpendicular line in the direction of greatest increase. Therefore, the gradient if normal to a level curve.

Exercise (9) "The directional derivative in the direction of $\mathbf{u}$ is given by $\nabla f \cdot \hat{\mathbf{u}}$." Why does this make sense?

Solution:

The gradient could be decomposed into a component in $\hat{\mathbf{u}}$ direction and a component normal to that. So the rate of change of the function in the $\hat{\mathbf{u}}$ direction is $\nabla f \cdot \hat{\mathbf{u}}$.

Exercise (10) If the gradient is zero, does that imply that you are at a max or a min? Why or why not?

Solution:

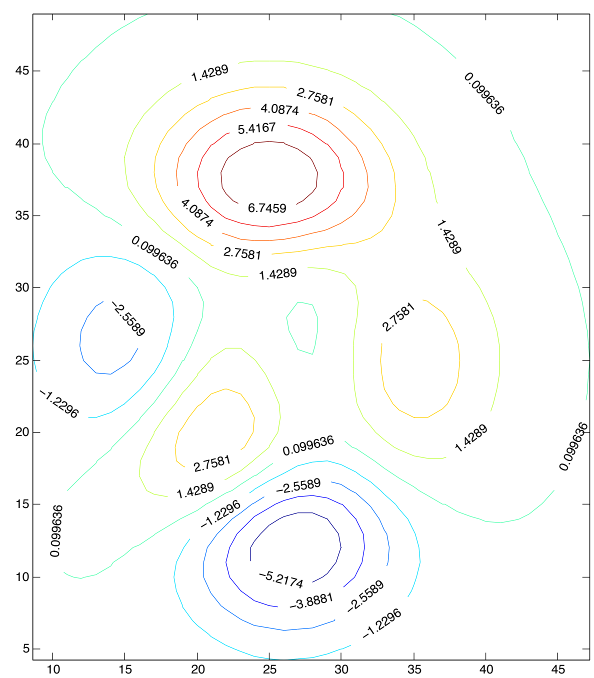
No, because you could be at a saddle point, for example.

Exercise (11) What does it mean for $\frac{\partial^2 f}{\partial x^2} > 0$? What about $\frac{\partial^2 f}{\partial x \partial y} > 0$?

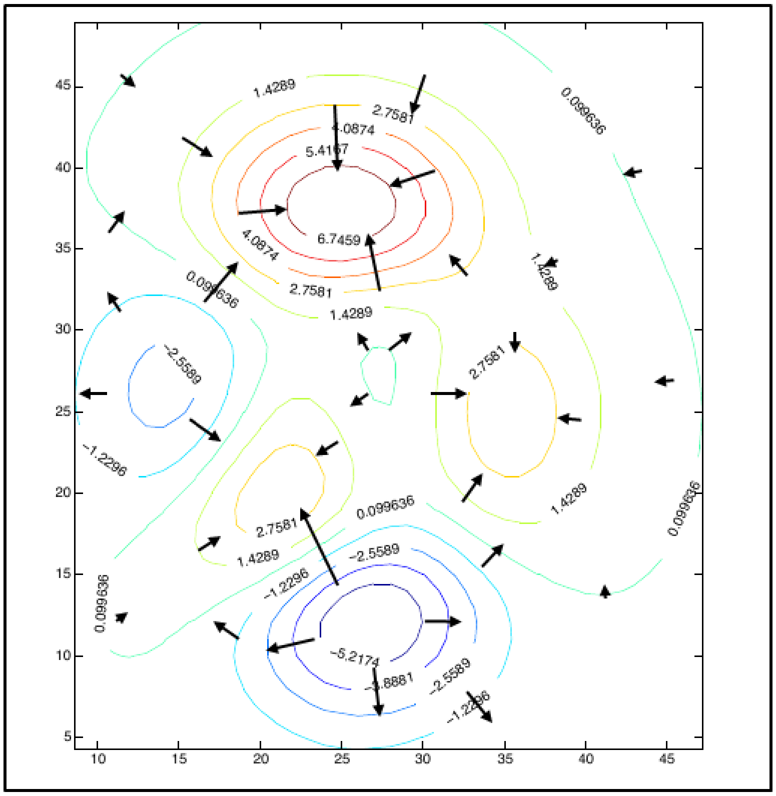
Solution:

Upward curvature in the x-direction; slope in x-direction increases with increasing y (and vice versa).

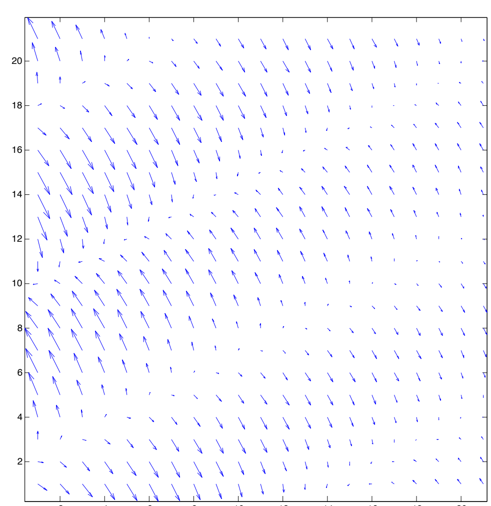
Exercise (12) The diagram below shows a contour plot for a function. Sketch in the gradient field.

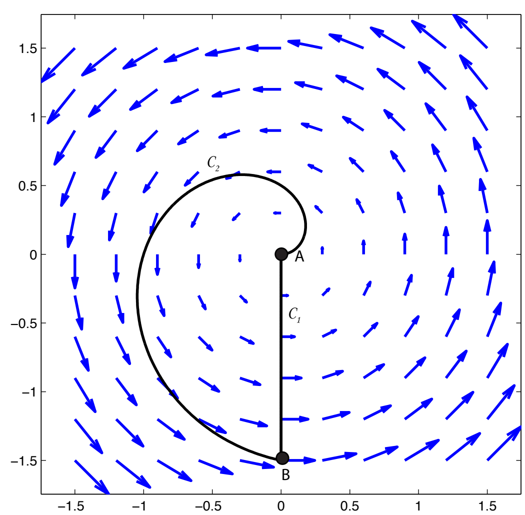


Solution:



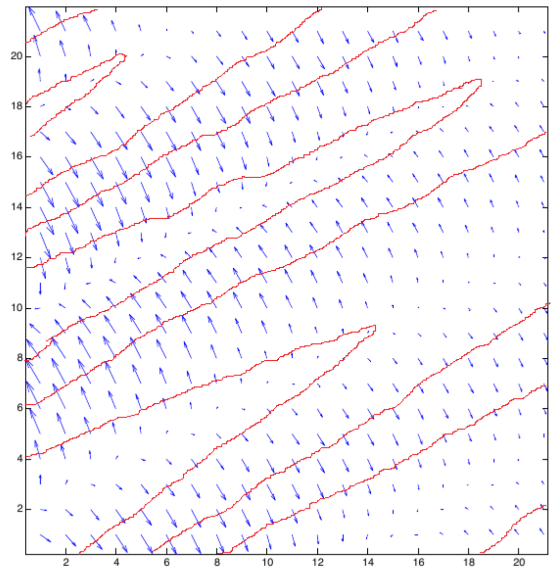
Exercise (13) The diagrams below show two vector fields. One is the gradient of a function of two variables; the other is not. Which one could be a gradient? Why? Sketch in the level curves for the one that works.



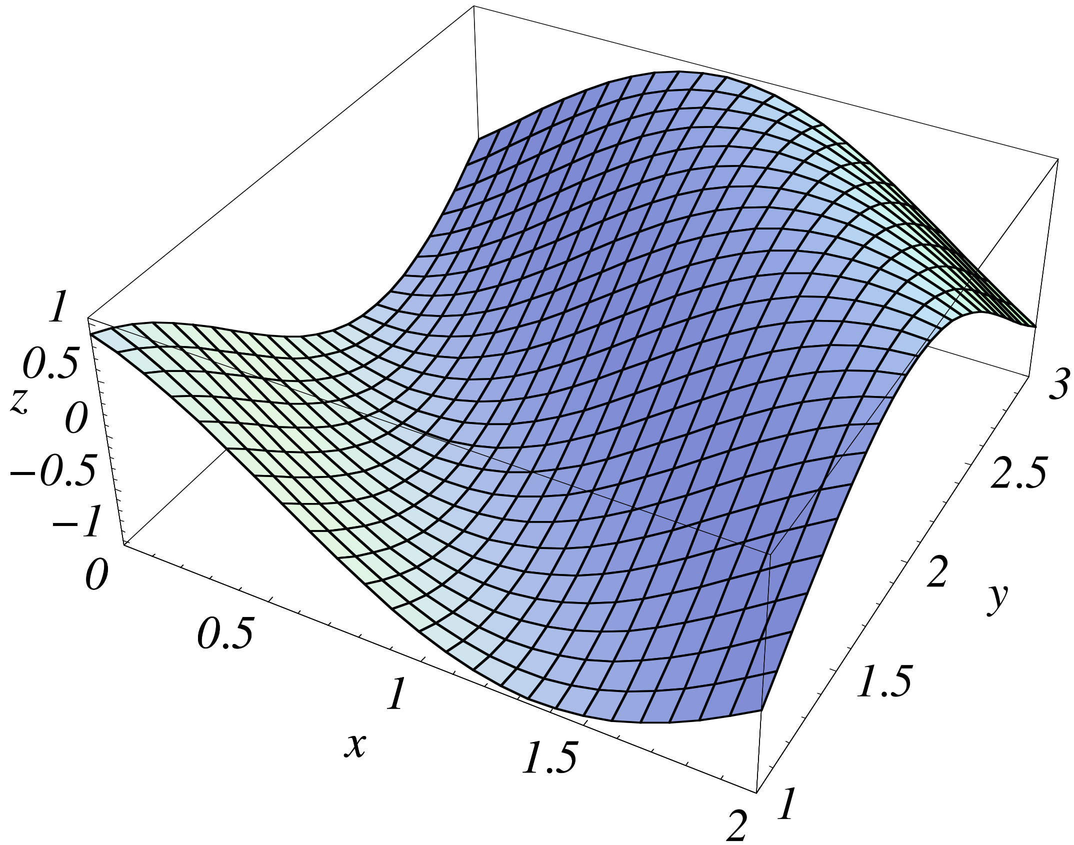


Solution:

The one on the left could be a gradient; the left one could not because a path of gradient ascent would go in a circle. See FIg. **??** for a sketchy sketch of level curves.



Exercise (14) A function f has the graph shown below. The grid lines on the graph are lines along which either the x- or the y- coordinate is held constant.



Determine the sign of each of the following:

(a) $f_x(1,2)$

(b) $f_y(1,2)$

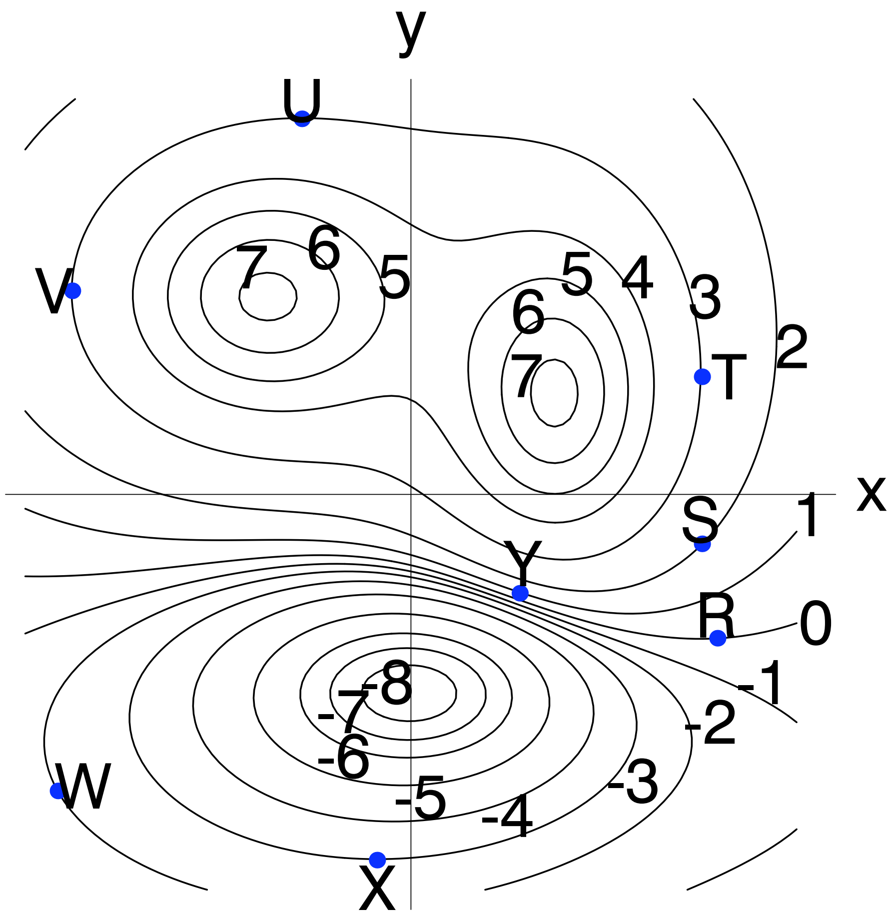
(c) $f_{xx}(1,2)$

(d) $f_{yy}(1,2)$

Solution:

(a)+; (b) +; (c) and (d) look pretty close to 0.

Exercise (15) The diagram below shows some level curves of a function $g(x, y)$. The numbers indicate the g-values of these level curves, and the letters indicate points on the level curves. Note that point Y is on the level curve $g = 1$ and point W is on the level curve $g = -2$.



(a) What are the signs of the partial derivatives $\frac{\partial g}{\partial x}$ and $\frac{\partial g}{\partial y}$ at each of the points marked? Why?

Solution:

Point	Sign of g_x	Sign of g_y
R	- or 0	+
S	-	+
T	-	0
U	0	-
V	+	0
W	-	-
X	0	-
Y	+	+

(b) At which of the points marked does the gradient vector ∇g have the greatest magnitude? Explain.

Solution:

At Y because the contours are closest together, indicating the greatest change in the function value per unit distance normal to the level curves.

(c) Let

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The directional derivative $D_{\hat{\mathbf{u}}} g$ is zero at exactly one of the points marked. Which point is it?

Solution:

Point S, where the contour line is parallel to $\mathbf{u}$.

4 *Readings, Videos, and Conceptual Questions - Optimization with Gradient Ascent [2 Hrs]*

At this link you will *also* find a set of readings and videos about gradient ascent (or descent - same thing, but for a sign!).

Read the text from Giordano and watch the videos. If you are interested, you can also read the stuff about conjugate gradient ascent, which is cool and leverages eigenstuff - but is optional!

Then answer the following questions:

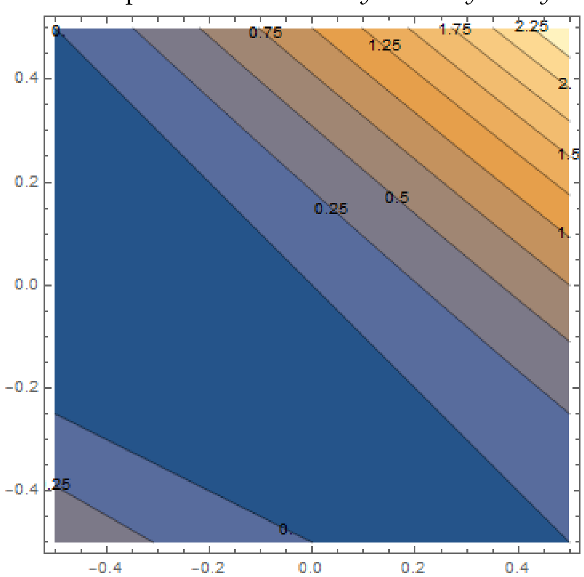
Exercise (16) At the origin, a given function has a gradient of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and a Hessian of

$$\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$$

Sketch a contour plot of the function in the vicinity of the origin.

Solution:

One example would be $x^2 + 2y^2 + x + y + 3xy$:



Important features are:

(a) Level curves get closer together as x and y increase (near the origin)

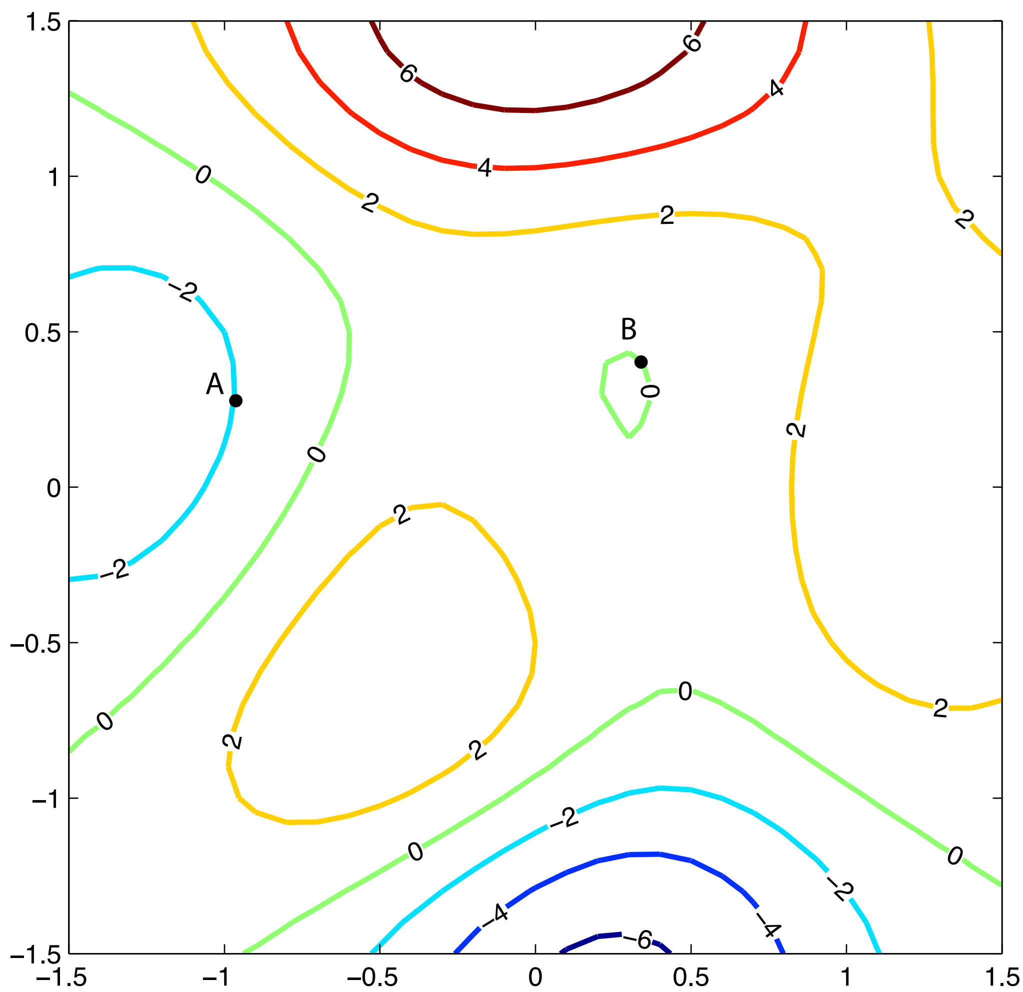
(b) Level curve at the origin is in direction $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

(c) Level curves near the origin tip upward (become more horizontal) as x and y increase

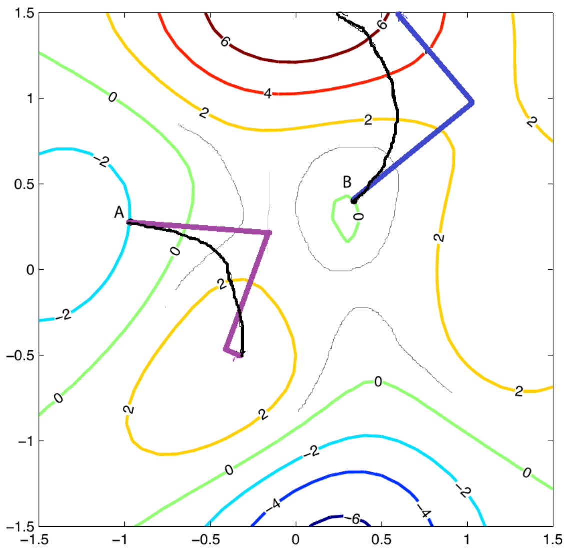
Exercise (17) The figure below shows a contour plot with two points marked (A and B). For both points,

(a) Draw the path that gradient ascent would use if the step size was small (following the approach in the first video).

(b) Draw the path that gradient ascent would follow if the algorithm is implemented as shown in the second video.



Solution:



The black path is the small step size ascent; the purple and blue paths are the walk-straight-until-you-start-to-go-downhill method. But it's a little hard to predict the path from point A given the large spacing of contours...other reasonable paths will be accepted.

5 *Gradient Ascent [2 Hrs]*

As you've just been learning, Gradient Ascent (or descent) is a technique to determine the maximum or minimum of a function of many variables by taking steps in the direction of the gradient (or negative gradient). If the height of a surface is described by $z = f(\mathbf{r})$, where $\mathbf{r}$ is the position vector in the plane, and we begin at $\mathbf{r}_0$, the points determined by Gradient Ascent are given by

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \lambda_i \nabla f(\mathbf{r}_i),\ i = 0, 1, 2, \ldots$$

where λ_i is the relative size of the step that we take in the direction of the gradient. There are various schemes for choosing these, and one of the simplest is to determine the next step with a simple proportionality

$$\lambda_{i+1} = \delta \lambda_i$$

where both δ and λ_0 are thoughtfully chosen for the problem at hand. We are going to develop a method and implementation to drive your NEATO on the floor of the classroom in a way that physically realizes the method of Steepest Ascent. First, we are going to introduce the mountain you will climb, and you will think through the steps involved.

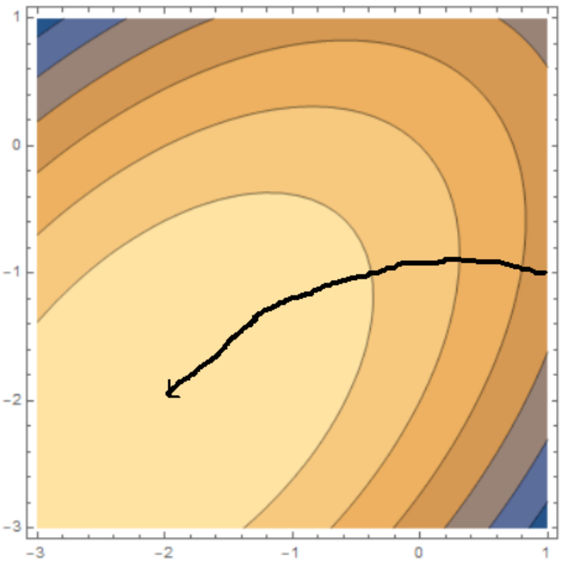
18. The mountain you will "climb" is defined as follows

$$f(x, y) = xy - x^2 - y^2 - 2x - 2y + 4$$

where x, y, and f are measured in feet. You are required to start at $(1, -1)$, and work your way to the top by method of steepest ascent.

(a) Visualize the contours of this function in MATHEMATICA on the domain $(-3, 1) \times (-3, 1)$, and print it out.

Solution:



(b) Draw the path of steepest ascent if we were moving continuously from a starting point at $(1, -1)$.

(c) Find the gradient of this function.

Solution:

$(-2 - 2x + y, -2 + x - 2y)$

(d) Assuming $\mathbf{r}_0 = (1, -1)$, what is the initial gradient at $\mathbf{r}_0$? What would be a reasonable choice for λ_0 so that $\mathbf{r}_1$ is not too far from the continuous path? Plot $\mathbf{r}_1$ on your contour plot.

Solution:

Initial gradient is (-5,1), with a magnitude of 5.1. To move about 0.5 feet, the initial multiplier λ_0 should be about 0.1 feet.

(e) Assuming you place your NEATO at $(1, -1)$ pointing along the y-axis, how much do you have to rotate it in order to align it with the gradient at $\mathbf{r}_0$? What would be a reasonable angular speed?

Solution:

1.37 rad (78.7 degrees) CCW. 0.5 rad/s would not exceed 0.3 m/s if rotated in place.

(f) Assuming that you are going to drive your NEATO at $0.1m/s$, how long would you drive in order to reach $\mathbf{r}_1$? (Careful with unit changes!)

Solution:

$t = \frac{|\mathbf{r}_1 - \mathbf{r}_0|}{V} \frac{1\text{m}}{3.28\text{ft}}$

(g) What is the gradient at $\mathbf{r}_1$? What value of δ should you use so that λ_1 and $\mathbf{r}_2$ are reasonable? Plot $\mathbf{r}_2$ on your contour plot.

(h) Assuming your NEATO is now at $\mathbf{r}_1$, how much do you have to rotate it in order to align it with the new gradient? What would be a reasonable angular speed?

(i) Assuming that you are going to drive your NEATO at $0.1 m/s$, how long would you drive in order to reach $\mathbf{r}_2$?