Module 9 Exercises | Joshua Shapiro | 19 June 2017

2. Why is this true? What is the minimum (unconstrained) value of f(x,y) = 3x + 4y?

- **For any linear, unconstrained problem, the minimum will always be -∞.**

3. Go back to your calc book and find an example of a "hard to differentiate" function.

- **Answer:** $\frac{4x^2+log(2^{tan(x^{sin(x)})})}{5x+cos(2^x)}$

4. What's an example of a function that's continuous but not differentiable? Consider the weird function f(x) where f(x) = 1 if x is rational and f(x) = 0 otherwise. Is this continuous? Differentiable?

- **The absolute value function is not differentiable.**
- **The weird function is not continuous and therefore not differentiable.**

5. Consider the function $f(x) = \frac{x}{\mu_1 - \lambda x} + \frac{1-x}{\mu_2 - \lambda(1-x)}$
   - Compute the derivative f'(x). Can you solve f'(x) = 0?

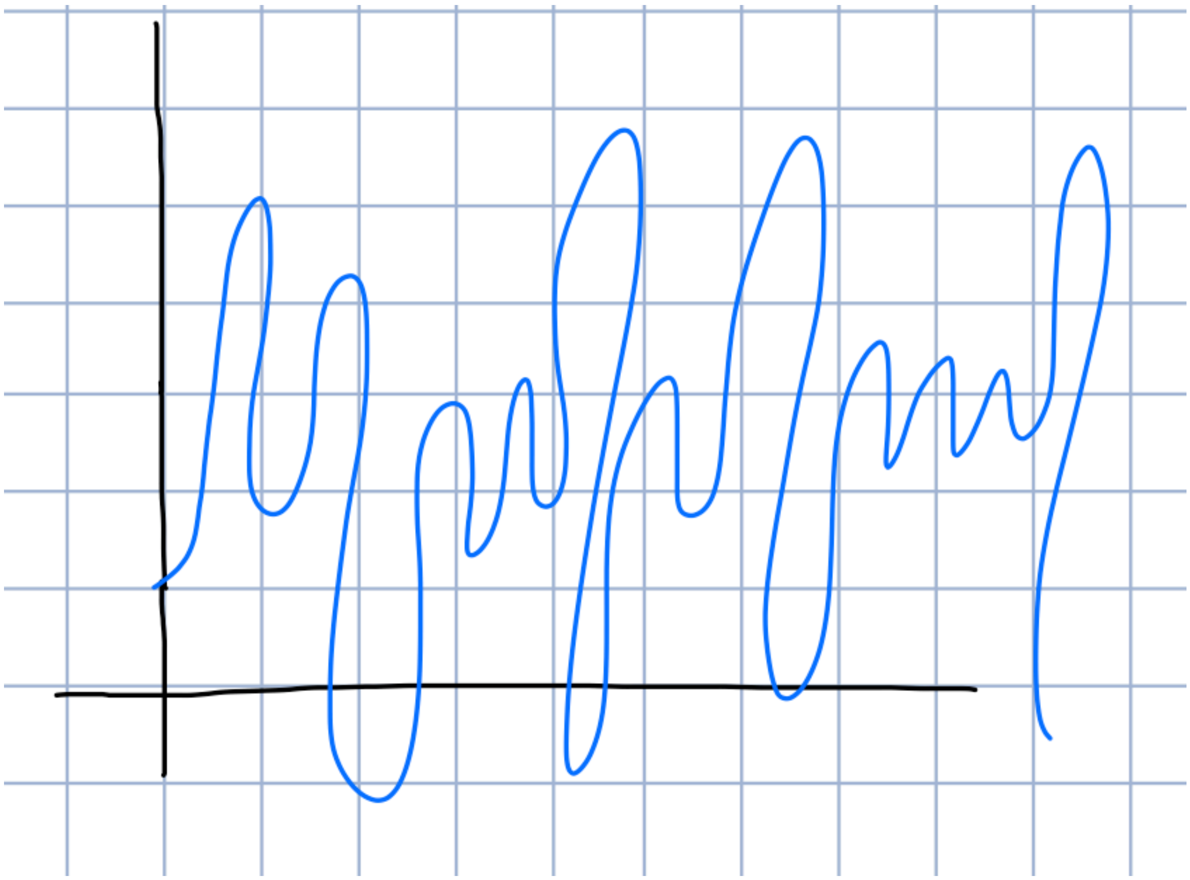
- **Derivative below:**
  - $f(x) = \frac{x}{\mu_1 - \lambda x} + \frac{1-x}{\mu_2 - \lambda(1-x)}$
  - $\implies f'(x) = \frac{\mu_1 - \lambda * x + \lambda * x}{mu_1 - \lambda * x)^2} + \frac{-\mu_2 + \lambda - \lambda * x - \lambda - \lambda * x}{(\mu_2 - \lambda + \lambda * x)^2}$
  - $\implies f'(x) = \frac{(\mu_1)(\mu_2 - \lambda + \lambda * x)^2 - (\mu_2)(\mu_1 - \lambda * x)^2}{(\mu_1 - \lambda * x)^2 (\mu_2 - \lambda + \lambda * x)^2}$
- **Solving for f'(x) = 0:**
  - $\frac{(\mu_1)(\mu_2 - \lambda + \lambda * x)^2 - (\mu_2)(\mu_1 - \lambda * x)^2}{(\mu_1 - \lambda * x)^2 (\mu_2 - \lambda + \lambda * x)^2} = 0$
  - $\implies (\mu_1)(\mu_2 + \lambda(x-1))^2 - (\mu_2)(\mu_1 - \lambda * x)^2 = 0$
  - $\implies \mu_1 \lambda^2 - \mu_1^2 \mu_2 - 2\mu_1 \lambda \mu_2 + \mu_1 \mu_2^2 + x(4\mu_1 \lambda \mu_2 - 2\mu_1 \lambda^2) + x^2(\mu_1 \lambda^2 - \lambda^2 \mu_2) = 0$
  - ...skipping a few steps...
  - $\implies x = \pm\sqrt{\frac{\mu_1 \mu_2 (\mu_1 - \lambda + \mu_2)^2}{\lambda^2 (\mu_1 - \mu_2)^2}} - \frac{4\mu_1 \lambda \mu_2 - 2\mu_1 \lambda^2}{2(\mu_1 \lambda^2 - \lambda^2 \mu_2)}$

6. CODE: Download and execute BracketSearch.java.
   - What is the running time in terms of M and N?
   - If we keep MN constant (e.g., MN = 24), what values of M and N produce best results?

- **The run time is O(MN)**
- **For the function in BracketSearch, M = 12 and N = 2 produced best results, with a minimum found at 2.5001 (actual minimum at 2.5).**

7. Draw an example of a function for which bracket-search fails miserably, that is, the true minimum is much lower than what's found by bracket search even for large M and N.

- **Any function with many peaks and valleys will cause this to fail.**

- 

8. What is the number of function evaluations in terms of M and N for the bracket-search algorithm?

- **There are (M+1)N + 1 function evaluations.**

9. What is the number of function evaluations in terms of M and N for the 2D bracket-search algorithm? How does this generalize to n dimensions?

- **There are** $((M+1)^2+1)N$ **function evaluations for the 2D bracket-search algorithm.**
- **There are** $((M+1)^n+1)N$ **function evaluations for the n-dimensional bracket-search algorithm.**

10. CODE: Add code to MultiBracketSearch.java to find the minimum of: $f(x_1, x_2) = (x_1 - 4.71)^2 + (x_2 - 3.2)^2 + 2(x_1 - 4.71)^2 * (x_2 - 3.2)^2$

- **With M = 6 and N = 4, x1 = 4.691358024691359, x2 = 3.2098765432098757, numFuncEvals=120**

11. CODE: Modify BracketSearch2.java to use the proportional-difference stopping condition.

- **With M = 10, N = 4, bestx = 4.72, bestf = 2.5001, prevBestf = 2.5080999999999998**

16. CODE: Download and execute GradientDemo.java
    - How many iterations does it take to get close to the optimum?
    - What is the effect of using a small α (e.g., $\alpha = 0.001$)?
    - In the method nextStep(), print out the current value of x, and the value of xf'(x) before the update.
    - Set $\alpha = 1$ Explain what you observe.
    - What happens when $\alpha = 10$?

- **It takes around 150 iterations to get close to the optimum.**

- **A smaller alpha requires more iterations to get to the optimum.**
- **Alpha is too large when set to 1 and jumps over the minimum, never able to find it. Its jumps are always consistent, so it alternates between the same values.**
- **When alpha is set to 10, it also is never able to find the minimum. However, the jumps are not consistent and always moves farther away from the minimum.**

17. CODE: Download GradientDemo2.java and examine the function being optimized.
    - Fill in the code for computing the derivative.
    - Try an initial value of x at 1.8. Does it converge?
    - Next, try an initial value of x at 5.8. What is the gradient at the point of convergence?

- **It converges at the global minimum with a value of x = 1.8.**
- **When x = 5.8 it only finds a local minimum**