EECS 16B Designing Information Devices and Systems II
Fall 2021 Discussion Worksheet

Discussion 11B

The following notes are useful for this discussion: Note 19.

1. Linear Approximation

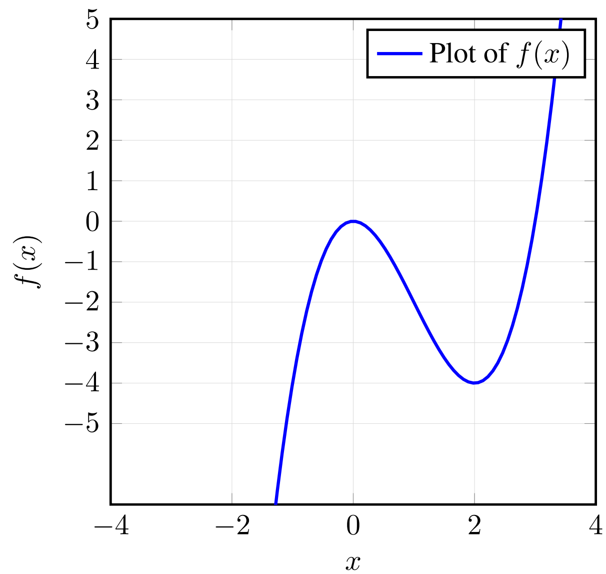
A common way to approximate a nonlinear function is to perform linearization near a point. In the case of a one-dimensional function $f(x)$, the linear approximation of $f(x)$ at a point $x_\star$ is given by

$$f(x) \approx f(x_\star) + f'(x_\star) \cdot (x - x_\star), \tag{1}$$

where $f'(x_\star) := \frac{\mathrm{d}f}{\mathrm{d}x}(x_\star)$ is the derivative of $f(x)$ at $x = x_\star$.

Keep in mind that wherever we see $x_\star$, this denotes a *constant value* or operating point.

(a) Suppose we have the single-variable function $f(x) = x^3 - 3x^2$. We can plot the function $f(x)$ as follows:



i. **Write the linear approximation of the function around an arbitrary point $x_\star$.**

ii. Using the expression above, **linearize the function around the point $x = 1.5$. Draw the linearization into the plot of part i).**

Now that we have this specific point's linearization, we understand how the function behaves around the point. Let's use this linearization to evaluate the function's approximation at $x = 1.7$ (based on our approximation at $x = 1.5$, we want to see how a $\delta = +0.2$ shift in the x value changes the corresponding $f(x)$ value). How does this approximation compare to the exact value of the function at $x = 1.7$?

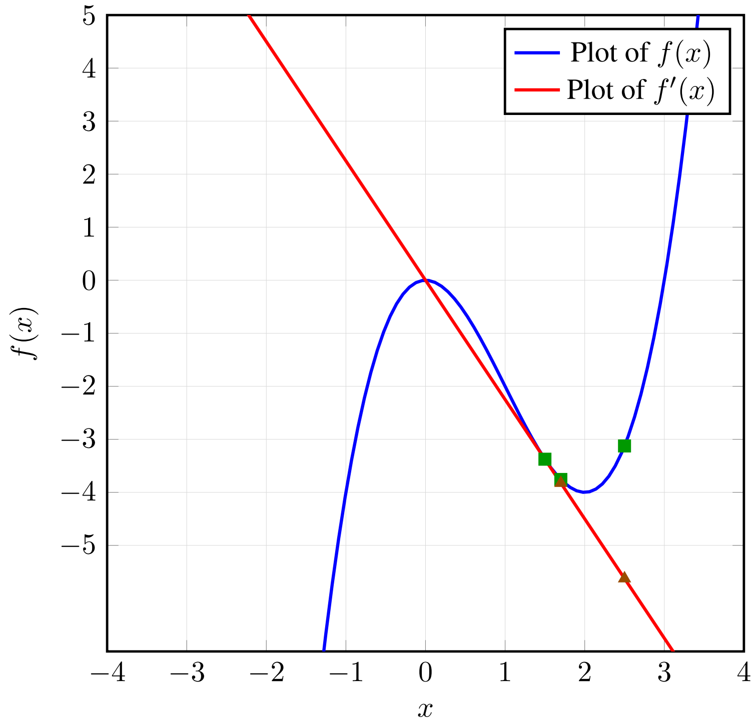
$$f(1.7) \approx -3.375 + (-2.25) \cdot (1.7 - 1.5) \tag{2}$$
$$\approx -3.375 - 0.45 \tag{3}$$
$$\approx -3.825 \tag{4}$$

Comparing to the exact value $f(1.7) = 1.7^3 - 3 \cdot 1.7^2 = -3.757$, we find that the difference is 0.068. Not too bad! What if we repeat with $\delta = 1$? To do so, we must use the approximation around $x = 1.5$ to compute $x = 2.5$, and compare to the exact value $f(2.5)$. How does our new approximation compare to the exact result?

$$f(2.5) \approx -3.375 + (-2.25) \cdot (2.5 - 1.5) \tag{5}$$
$$\approx -3.375 - 2.25 \tag{6}$$
$$\approx -5.625 \tag{7}$$

Comparing to the exact value $f(2.5) = 2.5^3 - 3 \cdot 2.5^2 = -3.125$, we find that the difference is much larger; the error jumped to 2.5! This is an error multiplication of $\frac{2.5}{0.068} \approx 37$, even though our δ only multiplied by 5. What happened?

Looking at the actual function, we see that the function has a significant curvature between our "anchor point" of $x_\star = 1.5$ and $x = 2.5$. Our linear model is unable to capture this curvature, and so we estimated $f(2.5)$ as if the function kept decreasing, as it did around $x = 1.5$ (where the slope was -2.25).



Now, we can extend this to higher dimensional functions. In the case of a two-dimensional function $f(x,y)$, the linear approximation of $f(x,y)$ at a point $(x_\star, y_\star)$ is given by

$$f(x,y) \approx f(x_\star, y_\star) + \frac{\partial f}{\partial x}(x_\star, y_\star)\cdot(x - x_\star) + \frac{\partial f}{\partial y}(x_\star, y_\star)\cdot(y - y_\star). \tag{8}$$

where $\frac{\partial f}{\partial x}(x_\star, y_\star)$ is the partial derivative of $f(x,y)$ with respect to x at the point $(x_\star, y_\star)$, and similarly for $\frac{\partial f}{\partial y}(x_\star, y_\star)$

(b) Now, let's see how we can find partial derivatives. When we are given a function $f(x,y)$, we calculate the partial derivative of f with respect to x by fixing y and taking the derivative with respect to x. **Given the function $f(x,y) = x^2 y$, find the partial derivatives $\frac{\partial f(x,y)}{\partial x}$ and $\frac{\partial f(x,y)}{\partial y}$.**

(c) **Write out the linear approximation of f near $(x_\star, y_\star)$.**

(d) We want to see if the approximation arising from linearization of this function is reasonable for a point close to our point of evaluation. **First, approximate** $f(x, y)$ **at the point** $(2.01, 3.01)$ **using** $(x_\star, y_\star) = (2, 3)$**. Next, compare the result to** $f(2.01, 3.01)$**.**

(e) We will now define the notion of a derivative as a function, and take a look at one possible representation of that function.

Given the representation of the derivative as a row-vector, describe a function that can take this representation, along with some column vector defining a change in direction, to return a scalar value (which is the change in the real-valued output). *Don't worry if this seems abstract for now, the next subpart will clarify.*

(f) Suppose we have now a scalar-valued function $f(\vec{x}, \vec{y})$, which takes in vector-valued arguments $\vec{x} \in \mathbb{R}^n$, $\vec{y} \in \mathbb{R}^k$ and outputs a scalar $\in \mathbb{R}$. That is, $f(\vec{x}, \vec{y})$ is $\mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}$. For this new model involving a vector-valued function, how can we adapt our previous linearization method?

One way to linearize the function f is to do it for every single element in $\vec{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^\top$ and $\vec{y} = \begin{bmatrix} y_1 & y_2 & \dots & y_k \end{bmatrix}^\top$. Then, when we are looking at x_i or y_j, we fix everything else as constant. This would give us the linear approximation

$$f(\vec{x}, \vec{y}) \approx f(\vec{x}_\star, \vec{y}_\star) + \sum_{i=1}^{n} \frac{\partial f(\vec{x}, \vec{y})}{\partial x_i} (x_i - x_{i,\star}) + \sum_{j=1}^{k} \frac{\partial f(\vec{x}, \vec{y})}{\partial y_j} (y_j - y_{j,\star}). \tag{9}$$

In order to simplify this equation, we can define the rows $D_{\vec{x}}$ and $D_{\vec{y}}$ as

$$D_{\vec{x}} f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_n} \end{bmatrix}, \tag{10}$$

$$D_{\vec{y}} f = \begin{bmatrix} \frac{\partial f}{\partial y_1} & \dots & \frac{\partial f}{\partial y_k} \end{bmatrix}. \tag{11}$$

Then, eq. (9) can be rewritten as

$$f(\vec{x}, \vec{y}) \approx f(\vec{x}_\star, \vec{y}_\star) + (D_{\vec{x}} f)\Big|_{(\vec{x}_\star, \vec{y}_\star)} \cdot (\vec{x} - \vec{x}_\star) + (D_{\vec{y}} f)\Big|_{(\vec{x}_\star, \vec{y}_\star)} \cdot (\vec{y} - \vec{y}_\star). \tag{12}$$

Assume that $n = k$ and we define the function $f(\vec{x}, \vec{y}) = \vec{x}^\top \vec{y} = \sum_{i=1}^{k} x_i y_i$. Find $D_{\vec{x}} f$ and $D_{\vec{y}} f$.

[Practice] Next, suppose $g(\vec{x}, \vec{y}) = x_1 x_2^2 y_1 + x_1 y_2^3 + x_2 x_1 y_2 y_1 + \frac{x_1^2}{x_2^3 y_1}$. Find $D_{\vec{x}} g$ and $D_{\vec{y}} g$

Hint: it can help to look at eq. (8), and match the terms in eq. (9) to that formulation.

(g) Following the above part, **find the linear approximation of** $f(\vec{x}, \vec{y})$ **near** $\vec{x}_\star = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ **and** $\vec{y}_\star = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$**.** **Recall that** $f(\vec{x}, \vec{y}) = \vec{x}^\top \vec{y} = \sum_{i=1}^{k} x_i y_i$**.**

These linearizations are important for us because we can do many easy computations using linear functions.

Contributors:

- Neelesh Ramachandran.
- Kuan-Yun Lee.

