

ARE201-Simon, Fall2015

CALCULUS1: TUE, SEP 8, 2015

PRINTED: SEPTEMBER 2, 2015

(LEC# 4)

CONTENTS

2. Univariate and Multivariate Differentiation	1
2.1. Univariate calculus: the derivative	1
2.2. The fundamental idea: linear approximations to nonlinear functions	2
2.3. Univariate Calculus: the differential	3
2.4. Multivariate calculus: functions from \mathbb{R}^n to \mathbb{R}	4
2.4.1. Partial Derivative	4
2.4.2. The Gradient	5
2.4.3. Crosspartial Derivative	5
2.4.4. The differential for functions from \mathbb{R}^n to \mathbb{R}	5

2. UNIVARIATE AND MULTIVARIATE DIFFERENTIATION

2.1. Univariate calculus: the derivative

Assume you all know how to calculate the derivative of a single variable function, i.e., given f , calculate $\frac{df(\cdot)}{dx}$, denoted also $f'(\cdot)$. Important to know the difference between $f'(\cdot)$, which is a function and $f'(x)$, which is a number, the function evaluated at a point.

I'll try to be careful to use this notation from now on: $g(\cdot)$ is a RULE, represents a function.

Since $f'(\cdot)$ is a function, like any other, it may have a derivative; if it does, call it $f''(\cdot)$. Do example $f(x) = x^2$.

Lots of standard kinds of functions you have to be able to differentiate in your sleep. Equivalent of being able to spell. Brainless activity.

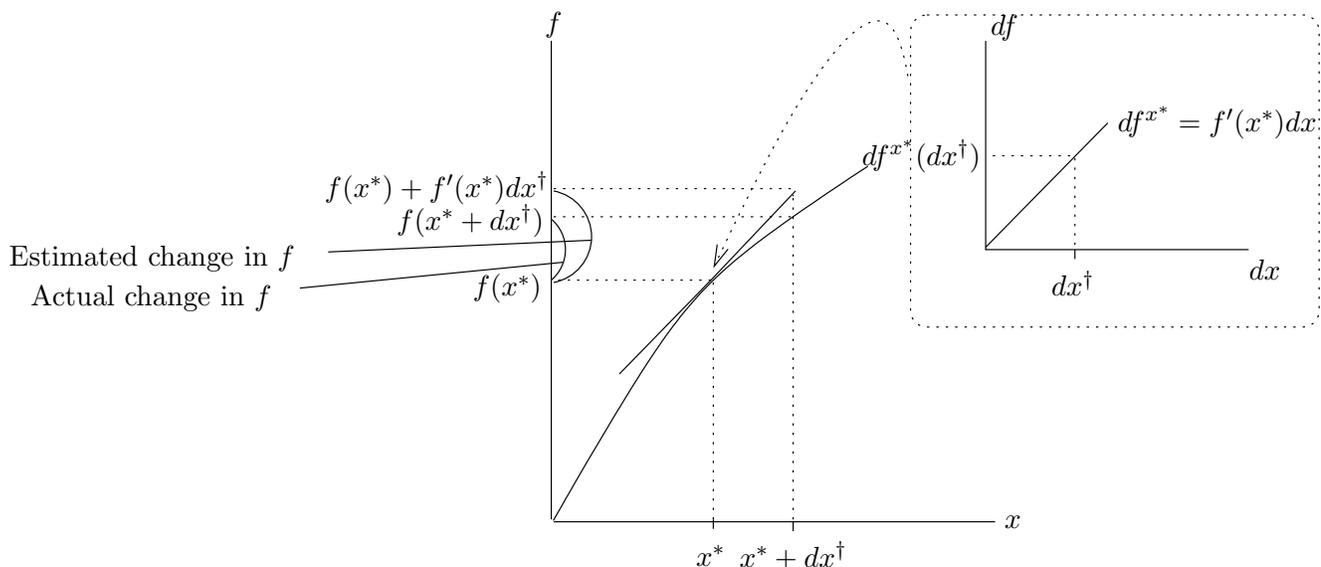


FIGURE 1. Linear Approximations of the change in a nonlinear function

2.2. The fundamental idea: linear approximations to nonlinear functions

The most important thing to keep in mind about calculus is:

(Calculus Mantra) When you do calculus, you are, ALWAYS, approximating a small change in a differentiable function by evaluating THE (unique) appropriate LINEAR function at the magnitude of the change. That is, for sufficiently small dx , $f(x + dx) - f(x) \approx f'(x)dx$.

$df^x(\cdot) = f'(x)(\cdot)$ is the linear function that approximates the difference between $f(x + \cdot)$ and $f(x)$.

Draw picture of the comparison: concave example.

- graph of $f(\cdot)$ with points x^* and $x^* + dx^*$.
- look at difference in the value of the function at the two points.
- look at the graph of the linear function $df = f'(x^*)dx$:
- note that the difference between $(f(x^* + dx^*) - f(x^*))$ and $f'(x^*)dx^*$ will be arbitrarily small provided that dx^* is sufficiently small.
- Explain the enormous practical importance of this.

When dx is bigger, don't do so well at approximation. Could do better if you threw in the second derivative, i.e.,

$$f(x + dx) - f(x) \quad \text{even better} \quad \approx \quad f'(x)dx + \frac{1}{2}f''(x)dx^2.$$

Note that in the above figure, $f''(x^*) < 0$. Note also that the linear approximation gives an overestimate of the change, so that adding the second term helps.

This is the germ of a crucial theorem called Taylor's theorem, that we'll see more of later.

2.3. Univariate Calculus: the differential

You've probably seen the expression $dy = f'(x)dx$, known as the *differential*, before. What sort of object is this? What's the relationship between the differential and the derivative? Recall from linear algebra about the relationship between linear functions from \mathbb{R}^1 to \mathbb{R}^1 and scalars; every such function is uniquely defined by a single scalar. E.g., I sometimes talked about the *function* p , when I really meant the linear function $f(\cdot)$ defined by $f(x) = px$. Well... The differential is to the derivative as the linear function is to the scalar/vector that defines it.

The relationship between the differential and the derivative is a particular example of the relationship between any linear function and the scalar/vector/matrix that defines it.

Defn: the differential of f at x is the linear function defined by the scalar $f'(x)$. In other words, the differential of f at x is the unique linear function whose coefficient (slope) is equal to the derivative of f at x .

Restate: The most important thing to keep in mind about calculus is: when you do calculus on f at x , you are, ALWAYS, approximating a small change in f , starting from x , by evaluating *the differential of f at x* at the magnitude of the change. This is the idea that underlies *all* of comparative statics.

That is, $dy = f'(x)dx$ is the value of the approximation to the change in f when you shift from x to $x + dx$.

Economic example: have $\pi(p)$, i.e., profits are a function of prices (assuming optimal input choices for prices p .) If p change is small enough, $d\pi^p \approx \pi'(p)dp$.

It is the ability to make computations like this—we call them comparative statics—that distinguish economists from art historians. Example:

- Let $f(x) = x^3 + 2x^2$
- set $\bar{x} = 1$, $dx = 0.1$
- then
 - $f(\bar{x}) = 3$,
 - $f'(\bar{x}) = 3\bar{x}^2 + 4\bar{x} = 7$,

- $f(\bar{x} + dx) = 1.1^3 + 2 * 1.21 = 3.751$.
- The true difference $f(\bar{x} + dx) - f(\bar{x})$ is .751.
- the estimated difference is $f'(\bar{x})dx = (3 \times 1^2 + 4 \times 1) \times dx = 0.70$, not bad.

A clarification about notation: some of you may have wondered about the relationship between ∂x and dx , and wondered when it's appropriate to use one or the other.

- at least in basic calculus, ∂x doesn't have any meaning by as a stand-alone symbol; it only appears as part of a ratio, e.g., $\frac{\partial f}{\partial x}$, representing the *limit* of ratios of finite changes.
- dx sometimes appears as part of a fraction, d.g., $\frac{df}{dx}$, but it also is used as an argument to the differential function, i.e., $f'(x)dx$, where it represents a small change.
- in summary, ∂x is never an element of the real line, but dx may be.

2.4. Multivariate calculus: functions from \mathbb{R}^n to \mathbb{R}

2.4.1. *Partial Derivative.* Cookbook approach: If $f : \mathbb{R}^n \rightarrow \mathbb{R}$, then $f_i(\cdot) \equiv \frac{\partial f(\cdot)}{\partial x_i}$ is the (usual) derivative of the single variable function you get by treating all other variables as constant.

Example:

$$\begin{aligned} f(\mathbf{x}) &= x_1^2 + 2x_1x_2 + x_2^2 \\ &= x_1^2 + 2x_1\alpha + \beta \\ f_1(\mathbf{x}) &= 2x_1 + 2\alpha + 0 \\ &= 2x_1 + 2x_2 \end{aligned}$$

A more graphical view of partial derivatives. Take a cross-section of the graph of the function along the i 'th axis, and look at the slope of the single-dimensional function you obtain in this way. This slope is the i 'th partial derivative.

Similarly, you could take a diagonal cross-section, and get a different one-dimensional slope. Generally, you get what is called a *directional derivative*. Partial derivatives are just special kinds of directional derivatives: in particular, they tell you the slopes you obtain when you take cross-sections along the various axes. We'll talk about directional derivatives much more in the next lecture.

2.4.2. *The Gradient.* Given a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, the gradient of f is just the function from \mathbb{R}^n to \mathbb{R}^n that maps each $\mathbf{x} \in \mathbb{R}^n$ to the vector of partial derivatives of f at \mathbf{x} . That is,

$$\nabla f(\cdot) = \begin{bmatrix} f_1(\cdot) \\ \vdots \\ f_i(\cdot) \\ \vdots \\ f_n(\cdot) \end{bmatrix}$$

$\nabla f(\cdot)$ is to $f : \mathbb{R}^n \rightarrow \mathbb{R}$. as $f'(\cdot)$ is to $f : \mathbb{R} \rightarrow \mathbb{R}$. In fact the words “slope of f ,” “gradient of f ” and “derivative of f ” are synonymous.

An important distinction to make is between “*the gradient of f* ” which is a *function*—written ∇f or $\nabla f(\cdot)$ —and *the “gradient of f at \mathbf{x} ,”*—i.e., $\nabla f(\mathbf{x})$ —which is a vector.

2.4.3. *Crosspartial Derivative.* A cross-partial derivative is just an entry in the matrix which is the derivative of the derivative. That is, take a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$. The i 'th partial derivative of this function $f_i(\cdot)$ is a function, like any other, and if the function is differentiable, it has derivatives:

$$f_{ij}(\cdot) \equiv \frac{\partial^2 f}{\partial x_i \partial x_j}(\cdot) \text{ is the } j\text{'th partial derivative of the function } f_i(\cdot) \equiv \frac{\partial f}{\partial x_i}(\cdot).$$

Thus, the Hessian of f is the matrix consisting of the crosspartials of f .

2.4.4. *The differential for functions from \mathbb{R}^n to \mathbb{R} .* Recall from earlier: the most important thing to keep in mind about calculus is that when you do calculus on f at x , you are, ALWAYS, approximating a small change in f , starting from x , by evaluating *the differential of f at x* at the magnitude of the change. We'll look at two numerical examples, then look at the picture.

Numerical example:

- $f(x, y) = x^2 + y^2$;
- $\nabla f(x, y) = (2x, 2y)$;
- Consider $(dx, dy) = (0.1, 0.1)$ and set $(x^*, y^*) = (100, 100)$
- approx $f(x^* + dx, y^* + dy) - f(x^*, y^*)$ by $\nabla f(x^*, y^*) \times (0.1, 0.1) = (200, 200) \times (0.1, 0.1) = 40$.
- Now evaluate at f at both $(x^*, y^*) = (100, 100)$ and $(x^* + dx, y^* + dy) = (100.1, 100.1)$, and compare: $f(x^*, y^*) = 20,000$; $f(x^* + dx, y^* + dy) = 20040.02$; Actual difference is 40.02.

Symbolic example:

- $y = f(\ell, k)$;
- $\nabla f(\ell, k) = (f_\ell(\ell, k), f_k(\ell, k))$.

- Fix (ℓ^*, k^*) and let $(\ell, k) = (\ell^*, k^*) + (d\ell, dk)$.
- Approximate the difference $f(\ell, k) - f(\ell^*, k^*)$ by the value of the differential, evaluated at $(d\ell, dk) = (\ell, k) - (\ell^*, k^*)$:

$$f(\ell, k) - f(\ell^*, k^*) = dy = f_\ell(\ell^*, k^*)d\ell + f_k(\ell^*, k^*)dk$$

Graphical example: In Fig. 2, we're doing exactly the same thing for functions from \mathbb{R}^n to \mathbb{R} as we did in Fig. 1 for functions from \mathbb{R} to \mathbb{R} . We evaluate the difference $f(\mathbf{x}^* + d\mathbf{x}^\dagger) - f(\mathbf{x}^*)$, where here $d\mathbf{x}^\dagger$ and \mathbf{x}^* are both vectors, by the value of the differential, evaluated at the difference, $d\mathbf{x}^\dagger$. What's the differential, in terms that we've used before? Recall that it is a linear function from \mathbb{R}^n to \mathbb{R} ; so it is characterized by a vector of coefficients. That is, $df^{\mathbf{x}^*} = \nabla f(\mathbf{x}^*) \cdot d\mathbf{x}$.

As Fig. 2 indicates, the differential here plays exactly the same role as it played for functions from \mathbb{R} to \mathbb{R} . In Fig. 2, the tangent plane lies below the function, at least a neighborhood of \mathbf{x}^* . Then, as we move from \mathbf{x}^* to $\mathbf{x}^* + d\mathbf{x}^\dagger$, the value of the function declines; but the height of the tangent plane declines even further. Applying the calculus mantra, we approximate the change in f , i.e., $f(\mathbf{x}^* + d\mathbf{x}^\dagger) - f(\mathbf{x}^*)$, by translating the tangent plane so that \mathbf{x}^* becomes the new origin, then approximate the change in f when you move from \mathbf{x}^* to $\mathbf{x}^* + d\mathbf{x}^\dagger$ by evaluating the height of this translated plane at the point $d\mathbf{x}^\dagger$. Notice from the top panel that the absolute value of $df^{\mathbf{x}^*}(d\mathbf{x}^\dagger)$ over-estimates the absolute magnitude of the change in the function.

The differential is sometimes referred to as the *total differential*. I'll hopefully never use this term, because it is too close to something that's different, i.e., *the total derivative*. The total derivative is related to (but not the same as) a directional derivative (i.e., you take the cross section of the function in a special direction). The total derivative is either a *number* $\frac{df(x)}{dx}$ or possibly a non-linear function $\frac{df(\cdot)}{dx}$. The total differential is the name of a linear *function*,. $df^{\mathbf{x}^*}(d\mathbf{x}) = \nabla f(\mathbf{x}^*)d\mathbf{x}$.

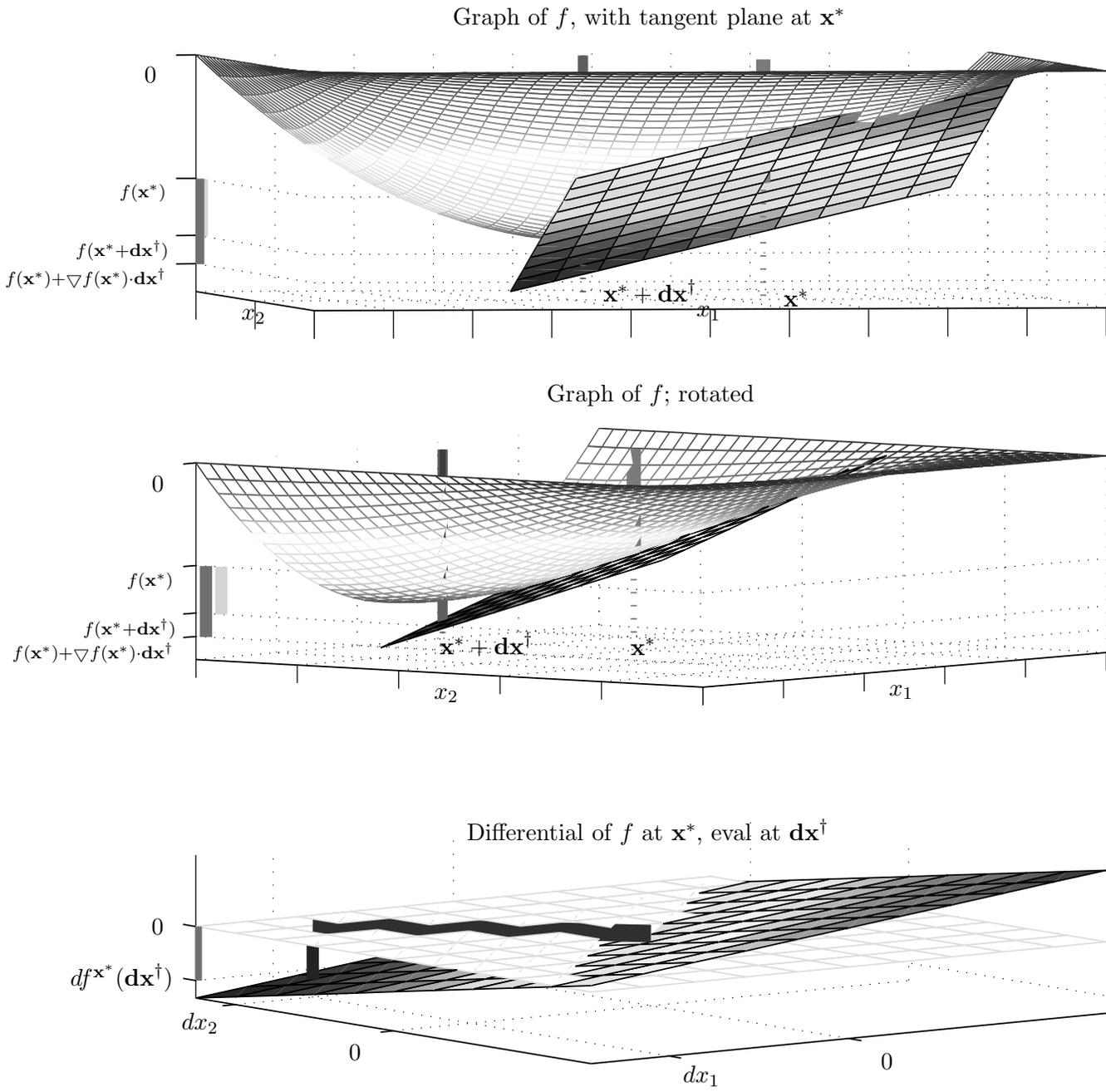


FIGURE 2. Linear Approximation of the change in a nonlinear multivariate function