Chapter 8
Multivariate Calculus

Isaac Barrow (1630-1677)

Augustin Louis Cauchy (1789–1857)

1

8.1 Multivariate Calculus: Partial Differentiation

- Now, y depends on several variables: $y = f(x_1, x_2, \ldots, x_n)$
- The derivative of y w.r.t. one of the variables –while the other variables are held constant- is called a *partial derivative*.

$$y = f(x_1, x_2, \ldots, x_n)$$

$$\lim_{\Delta x_1 \to 0} \frac{\Delta y}{\Delta x_1} = \lim_{\Delta x_1 \to 0} \frac{f(x_1 + \Delta x_1, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{\Delta x_1}$$

$$\equiv \frac{\partial y}{\partial x_1} \equiv f_1 \qquad \text{(partial derivative w.r.t. } x_1\text{)}$$

In general, $$\lim_{\Delta x_i \to 0} \frac{\Delta y}{\Delta x_i} \equiv \frac{\partial y}{\partial x_i} \equiv f_i, \quad i = 1 \ldots n$$

2

8.1 Partial Differentiation: Example

Cobb - Douglas production function : $Q = AK^{\alpha} L^{\beta}$

$(A = 96; \alpha + \beta = 1)$ $\qquad Q = 96K^{0.3} L^{0.7}$

$$MPP_K = \frac{\partial Q}{\partial K} = (0.3)96K^{-.7}L^{0.7} = 28.8K^{-.7}L^{0.7}$$

$$MPP_L = \frac{\partial Q}{\partial L} = (0.7)96K^{0.3}L^{-0.3} = 67.2K^{0.3}L^{-0.3}$$

We collect the first derivatives in a vector, ∇Q.

$$\nabla Q = \begin{bmatrix} \frac{dQ}{dx} \\ \frac{dQ}{dx} \end{bmatrix} = \begin{bmatrix} 28.8\,K^{-0.7}L^{0.7} \\ 67.2\,K^{0.3}L^{-0.3} \end{bmatrix}$$

Note: The first derivative of a scalar function w.r.t a vector is called the *gradient*.

3

8.1 Partial differentiation: Market Model

$$Q_d = a - bP \qquad (a, b > 0)$$

$$Q_s = -c + dP \qquad (c, d > 0)$$

$$\begin{bmatrix} 1 & b \\ 1 & -d \end{bmatrix} \begin{bmatrix} Q \\ P \end{bmatrix} = \begin{bmatrix} a \\ -c \end{bmatrix}$$

$$\frac{1}{-(b+d)} \begin{bmatrix} -d & -b \\ -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ -c \end{bmatrix} = \begin{bmatrix} Q^* \\ P^* \end{bmatrix}$$

$$Q^* = \frac{ad - bc}{b + d} \qquad P^* = \frac{a + c}{b + d}$$

$$\frac{\partial Q^*}{\partial a} = \frac{d}{b+d} > 0 \qquad \frac{\partial Q^*}{\partial c} = \frac{-b}{b+d} < 0$$

$$\frac{\partial P^*}{\partial a} = \frac{1}{b+d} > 0 \qquad \frac{\partial P^*}{\partial c} = \frac{1}{b+d} > 0$$

4

8.1 Partial differentiation: Market Model

- Using linear algebra, we have:

$$\begin{bmatrix} Q^* \\ P^* \end{bmatrix} = \begin{bmatrix} \dfrac{d}{(b+d)} & \dfrac{b}{(b+d)} \\ \dfrac{1}{(b+d)} & \dfrac{-1}{(b+d)} \end{bmatrix} \begin{bmatrix} a \\ -c \end{bmatrix}$$

$$x^* = A^{-1}d$$

$$\partial x^* / \partial d = A^{-1}$$

$$\frac{\partial x^*}{\partial d} = \begin{bmatrix} \dfrac{\partial Q^*}{\partial a} & \dfrac{\partial Q^*}{\partial c} \\ \dfrac{\partial P^*}{\partial a} & \dfrac{\partial P^*}{\partial c} \end{bmatrix} = \begin{bmatrix} \dfrac{d}{b+d} & \dfrac{-b}{b+d} \\ \dfrac{1}{b+d} & \dfrac{1}{b+d} \end{bmatrix}$$

<u>Note</u>: We call the matrix of first partial derivatives with respect to a vector the *Jacobian*, **J**.

5

8.1 Partial Differentiation: Likelihood

- In the usual estimation problem in Classical Linear Model (CLM), the unknowns are the parameters (typical in the CLM, β and σ^2). We treat the data (x_t and y_t) as (conditionally) known numbers. Assuming normality for the error term, ε_t, the (log) likelihood function is:

$$\log L = -\frac{T}{2}\ln(2\pi\sigma^2)^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\varepsilon_t^2 = -\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t - x_{1,t}\beta_1 - x_{2,t}\beta_2)^2$$

1st partial derivatives:

$$\frac{\partial \ln L}{\partial \beta_1} = -\frac{1}{2\sigma^2}\sum_{t=1}^{T} 2(y_t - x_{1,t}\beta_1 - x_{2,t}\beta_2)(-x_{1,t}) = \frac{1}{\sigma^2}\sum_{t=1}^{T}(y_t x_{1,t} - x_{1,t}^2\beta_1 - x_{2,t}x_{1,t}\beta_2)$$

$$\frac{\partial \ln L}{\partial \beta_2} = \frac{1}{\sigma^2}\sum_{t=1}^{T}(y_t x_{2,t} - x_{1,t}x_{2,t}\beta_1 - x_{2,t}^2\beta_2)$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{t=1}^{T}\varepsilon_t^2$$

$$J = \left[\frac{1}{\sigma^2}\sum_{t=1}^{T}(y_t x_{1,t} - x_{1,t}^2\beta_1 - x_{2,t}x_{1,t}\beta_2) \quad \frac{1}{\sigma^2}\sum_{t=1}^{T}(y_t x_{2,t} - x_{1,t}x_{2,t}\beta_1 - x_{2,t}^2\beta_2) \quad -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{t=1}^{T}\varepsilon_t^2\right]$$

6

8.1 The Jacobian

- The *Jacobian* is the matrix of first partial derivatives at the point **x** (with respect to a vector):

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

<u>Notation</u>: **J** or Df_x. For the one equation case (a scalar function), **J** is a row vector and it's usually called *gradient* or *gradient vector at* **x.** It is usually written as a column vector as $\nabla f(\mathrm{x})$ and also called the *gradient* or *gradient vector at* **x.**

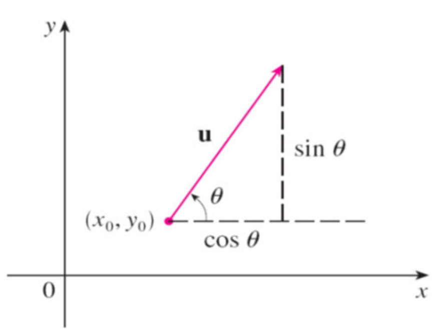
- A vector is characterized by its length and direction. To emphasize the direction, the length, ***h***, can be standardized, say $\|\boldsymbol{h}\| = 1$. The direction is studied with *directional derivatives.*

7

8.1 Directional Derivatives

- We can think that the partial derivatives of $z = f(x, y, w, \ldots)$ represent the rates of changes of z in the $x, y, \ldots$.

- Suppose that we now wish to find the rate of change of z at (x_0, y_0) in the direction of an arbitrary unit vector $\mathbf{u} = \langle a, b \rangle$.



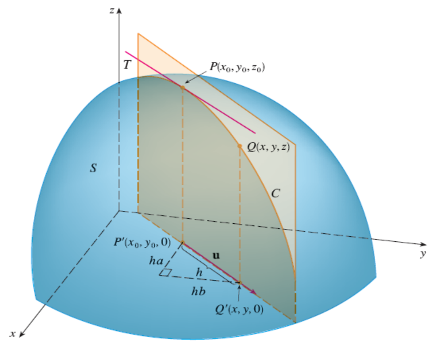
8

8.1 Directional Derivatives

- Consider the surface S with equation $z = f(x, y)$ [the graph of f] and we let $z_0 = f(x_0, y_0) \Rightarrow$ The point $P(x_0, y_0, z_0)$ lies on S.

- The vertical plane that passes through P in the direction of **u** intersects S in a curve C.

- The slope of the tangent line T to C at the point P is the rate of change of z in the direction of **u**.



8.1 Directional Derivatives

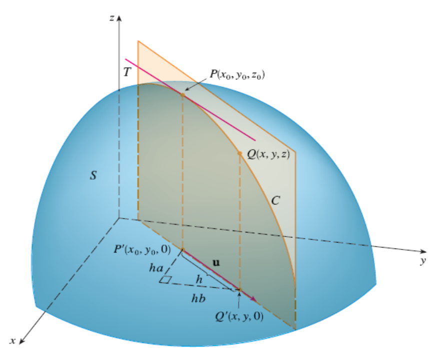
- Now, let $Q(x, y, z)$ be another point on C.
  - P', Q' be the projections of P, Q on the xy-plane.
  - The vector $\overrightarrow{P'Q'}$ is parallel to **u**.
  - $\overrightarrow{P'Q'} = h\,\mathbf{u} = \langle h\,a, h\,b\rangle$, for some scalar h.

- Then,

$$x - x_0 = ha \;\Rightarrow\; x = x_0 + ha$$
$$y - y_0 = hb \;\Rightarrow\; y = y_0 + hb$$

$$\frac{\Delta z}{h} = \frac{f(x_o + ha, y_o + hb) - f(x_0, y_0)}{h}$$

If we take the limit as $h \to 0$, we get the rate of change of z (w. r. to distance) in the direction of **u**.



8.1 Directional Derivatives

- The *directional derivative* of f at (x_0, y_0) in the direction of a unit vector $\mathbf{u} = \langle a, b \rangle$ is:

$$D_u f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_o + ha, y_o + hb) - f(x_0, y_0)}{h}$$

if the limit exists.

- Special cases:

- If $\mathbf{u} = \mathbf{i} = \langle 1, 0 \rangle$, then $D_i f = f_x$.
- If $\mathbf{u} = \mathbf{j} = \langle 0, 1 \rangle$, then $D_j f = f_y$.

That is, the partial derivatives of f with respect to x and y are just special cases of the directional derivative.

11

8.1 The Jacobian Determinant

- The Jacobian determinant, $|\mathbf{J}|$, at a point $\mathbf{x}$ gives information about the behavior of $F(.)$ near $\mathbf{x}$. For instance, the continuously differentiable function F is invertible near a point $\mathbf{x} \in R^n$ if $|\mathbf{J}| \neq 0$.

- Use $|\mathbf{J}|$ to test the existence of functional dependence between functions. If $|\mathbf{J}| = 0 \Rightarrow$ functional dependence, that is, a solution to a system of equations does not exist.
- Not limited to linear functions.
- For the 2x2 case:

$$|J| = \begin{vmatrix} \partial y_1 / \partial x_1 & \partial y_1 / \partial x_2 \\ \partial y_2 / \partial x_1 & \partial y_2 / \partial x_2 \end{vmatrix}$$

Carl Jacobi (1804 – 1851, Germany)

12

8.1 Cross partial derivatives

- The partial derivative is also a function of $\boldsymbol{x}$: $f'(\boldsymbol{x}) = g(x_1, x_2, \ldots, x_n)$

- If the n partial derivatives are continuous functions at point $\boldsymbol{x}$, we say that f is *continuously differentiable* at $\boldsymbol{x}$.

- If the n partial derivatives are themselves differentiable on an open set S $\in R^n$, we can compute their partial derivatives. For example:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$$

- The result of this differentiation is known as the *cross partial derivative* of f with respect to x_i and x_j. It is usually denoted as f_{ij}.

- When $i=j$, cross partial derivatives becomes the second-order derivative, denoted as f_{ii}. The matrix of all second derivatives is the Hessian.

13

8.1 Cross partial derivatives: Greeks

- We want to know how the BS Δ changes as maturity approaches.

Recall:

$\Delta = \frac{dC_t}{dS_t} = N(d1)$ & d1 = $[\ln(S_t/K) + (i + \sigma^2/2)(T - t)]/(\sigma\sqrt{T-t})$.

Then,

$$\frac{d^2 C_t}{dS_t dt} = \frac{d\Delta}{dt} = \frac{dN(d1)}{dt} = N'(d1) * \frac{d(d1)}{dt}$$

Using $N'(d1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{d1^2}{2}}$ &

$$\frac{d(d1)}{dt} = -\frac{\left(i + \frac{\sigma^2}{2}\right)}{2\sigma\sqrt{T-t}} + \frac{ln\left(\frac{S}{K}\right)}{2\sigma(T-t)^{3/2}}$$

we get:

$$\frac{d\Delta}{dt} = \frac{1}{\sqrt{2\pi}} e^{-\frac{d1^2}{2}} * \left\{-\frac{\left(i + \frac{\sigma^2}{2}\right)}{2\sigma\sqrt{T-t}} + \frac{ln\left(\frac{S}{K}\right)}{2\sigma(T-t)^{3/2}}\right\}$$

14

8.1 Cross partial derivatives: Greeks

- If the option is at-the-money (S_t = K), then

$$\frac{d\Delta}{dt} = \frac{1}{\sqrt{2\pi}} e^{-\frac{d1^2}{2}} * \left\{ -\frac{\left(i+\frac{\sigma^2}{2}\right)}{2\sigma\sqrt{T-t}} \right\} < 0$$

That is, as time goes by, delta decreases.

Note: Same qualitative result if the call option is out-of-the-money (S_t < K).

15

8.1 Cross partial derivatives: Hessian

- The matrix of all second derivatives is called the *Hessian*, usually denoted by **H**. For example:

$$H = \begin{bmatrix} \partial^2 f / \partial x_1^2 & \partial^2 f / (\partial x_1 \partial x_2) \\ \partial^2 f / (\partial x_2 \partial x_1) & \partial^2 f / \partial x_2^2 \end{bmatrix}$$

Example: Cobb-Douglas function, Q = $AK^{\alpha}L^{\beta}$

$$Q_K = \alpha A K^{\alpha-1} L^{\beta}$$

$$Q_L = \beta A K^{\alpha} L^{\beta-1}$$

$$H = \begin{bmatrix} -\alpha(1-\alpha) A K^{\alpha-2} L^{\beta} & \beta\alpha A K^{\alpha-1} L^{\beta-1} \\ \alpha\beta A K^{\alpha-1} L^{\beta-1} & -\beta(1-\beta) A K^{\alpha} L^{\beta-2} \end{bmatrix}$$

Note: $f_{ij} = f_{ji}$. This is a general result (Young's Theorem). Then, **H** is a symmetric matrix. **H** plays a very important role in optimization.

16

8.1 Cross partial derivatives: Hessian - Example

- We want to calculate **H** for a function, using econometrics notation, of β_1 and β_2 and σ^2 (we treat x_t and y_t as constants, along with σ^2). This (log) function is:

$$\log L = f(y_1, y_2, ..., y_T \mid \beta, \sigma^2) = -\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_t - x_{1,t}'\beta_1 - x_{2,t}'\beta_2)^2$$

1st derivatives:

$$f_x = \frac{\partial \ln L}{\partial \beta_1} = -\frac{1}{2\sigma^2}\sum_{t=1}^{T} 2(y_t - x_{1,t}\beta_1 - x_{2,t}\beta_2)(-x_{1,t}) = \frac{1}{\sigma^2}\sum_{t=1}^{T}(y_t x_{1,t} - x_{1,t}^2\beta_1 - x_{2,t}x_{1,t}\beta_2)$$

$$f_y = \frac{\partial \ln L}{\partial \beta_2} = \frac{1}{\sigma^2}\sum_{t=1}^{T}(y_t x_{2,t} - x_{1,t}x_{2,t}\beta_1 - x_{2,t}^2\beta_2)$$

2nd derivatives and cross derivatives:

$$f_{xx} = \frac{\partial^2 \ln L}{\partial \beta_1^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{T} x_{1,t}^2 < 0$$

$$f_{xy} = \frac{\partial^2 \ln L}{\partial \beta_1 \partial \beta_2} = -\frac{1}{\sigma^2}\sum_{t=1}^{T} x_{2,t}x_{1,t}$$

$$f_{yy} = \frac{\partial^2 \ln L}{\partial \beta_2^2} = -\frac{1}{\sigma^2}\sum_{t=1}^{T} x_{2,t}^2 < 0$$

17

8.1 Cross partial derivatives: Hessian - Example

- Then:

$$f_{xx} = -\frac{1}{\sigma^2}\sum_{t=1}^{T} x_{1,t}^2 < 0$$

$$f_{xy} = -\frac{1}{\sigma^2}\sum_{t=1}^{T} x_{2,t}x_{1,t}$$

$$f_{yy} = -\frac{1}{\sigma^2}\sum_{t=1}^{T} x_{2,t}^2 < 0$$

$$H = \begin{bmatrix} -\frac{1}{\sigma^2}\sum_{t=1}^{T}(x_{1,t}^2) & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(x_{2,t}x_{1,t}) \\ -\frac{1}{\sigma^2}\sum_{t=1}^{T}(x_{1,t}x_{2,t}) & -\frac{1}{\sigma^2}\sum_{t=1}^{T}(x_{2,t}^2) \end{bmatrix} = -\frac{1}{\sigma^2}\sum_{t=1}^{T}\begin{bmatrix} x_{1,t}^2 & x_{2,t}x_{1,t} \\ x_{1,t}x_{2,t} & x_{2,t}^2 \end{bmatrix} = -\frac{X'X}{\sigma^2}$$

<u>Note</u>: **H** plays an important role in maximum likelihood estimation. Its (negative expected) inverse is used to calculate SE.

18

8.2 Differentials

Problem: What if no explicit reduced-form solution exists because of the general form of the model?

Example: In the macro model, what is $\partial Y / \partial T$ when

$$Y = C(Y, T_0) + I_0 + G_0 \ ?$$

T_0 can affect C direct and indirectly through Y, violating the partial derivative assumption.

Solution: Use differentials! Recall that we thought of differentials as a (1^{st}-order) approximation to a change in f(x): $df(x) = \Delta x\, f'(x)$

- Find the derivatives directly from the original equations in the model.
- Take the total differential, adding all the effects (indirect and direct).
- The partial derivatives become the parameters in the sum.

19

8.2.1 Differentials and derivatives

- Recall that we thought of differentials as a (1^{st}-order) approximation to a change in f(x): $df(x) = \Delta x\, f'(x)$.
- Total derivatives measure the total change in y from the *direct* and *indirect* affects of a change in x_i.
- The symbols dy and dx are called the *differentials* of y & x, respectively.
- A differential describes the change in y that results for a specific and *not necessarily* small change in x from any starting value of x in the domain of the function $y = f(x)$.
- The derivative (dy/dx) is the quotient of two differentials: dy & dx.
- $f'(x)dx$ is a first-order approximation of dy:

$$y = f(x) \quad \Rightarrow \quad dy = f'(x)\, dx$$

20

8.2.2 Differentials and point elasticity

- Let $Q_d = f(P)$ (explicit-function general-form demand equation)
- Find the elasticity of demand with respect to price. We use and manipulate differentials.

$$\varepsilon_d \equiv \frac{\%\Delta Q_d}{\%\Delta P} = \frac{(dQ_d)/Q_d}{(dP)/P} = \frac{\left(dQ_d/dP\right)}{Q_d/P} = \frac{\text{marginal function}}{\text{average function}}$$

$$elastic \quad if\ |\varepsilon_d| > 1,\ inelastic \quad if\ |\varepsilon_d| < 1$$

21

8.3 Total Differentials

- Extending the concept of differential to smooth continuous functions with two or more variables
- Let $y = f(x_1, x_2)$ Find total differential dy

$$dy = \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2 \qquad \Rightarrow \quad dy = f_1 dx_1 + f_2 dx_2$$

- Chain Rule derivation:

Find dz/dx_1, where $z = f(y)$ and $y = g(x_1, x_2)$.

Algorithm : Substitute the total differential of y into that of z and divide through by dx_1 assuming $dx_2 = 0$

1) $dz = \frac{dz}{dy} dy$

2) $dy = \frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2$

3) $dz = \frac{dz}{dy}\left(\frac{\partial y}{\partial x_1} dx_1 + \frac{\partial y}{\partial x_2} dx_2\right)$

4) $\frac{dz}{dx_1}\Big|_{dx_2=0} = \frac{dz}{dy}\frac{\partial y}{\partial x_1}$

22

8.3 Total Differentials - Example

- Let U be a utility function: U = U ($x_1, x_2, \ldots, x_n$)
- Differentiation of U with respect to x_i
- $\partial U / \partial x_i$ is the marginal utility of the good x_i
- dx_i is the change in consumption of good x_i.

$$dU = \frac{\partial U}{\partial x_1} dx_1 + \frac{\partial U}{\partial x_2} dx_2 + \ldots + \frac{\partial U}{\partial x_n} dx_n$$

• dU equals the sum of the marginal changes in the consumption of each good and service in the consumption function.
• To find total derivative wrt to x_1 divide through by the differential dx_1 (partial total derivative):

$$\frac{dU}{dx_1} = \frac{\partial U}{\partial x_1} + \frac{\partial U}{\partial x_2}\frac{dx_2}{dx_1} + \ldots + \frac{\partial U}{\partial x_n}\frac{dx_n}{dx_1}$$

23

8.3 Rules of Differentials (same as derivatives)

Let k is a constant function; u = $u(x_1)$; v = $v(x_2)$

- 1. $dk = 0$ (constant-function rule)
- 2. $d(cu^n) = cn\, u^{n-1}\, du$ (power-function rule)
- 3. $d(u \pm v) = du \pm dv$ (sum-difference rule)
- 4. $d(uv) = v\, du + u\, dv$ (product rule)
- 5. $d\left(\frac{u}{v}\right) = \frac{vdu - udv}{v^2}$ (quotient rule)
- 7. $d(uvw) = vw\, du + uw\, dv + uv\, dw$

$$d(u \pm v \pm w) = du \pm dv \pm dw$$

24

8.3 Example:

Find the total differential (dz) of the function

1) $z = \dfrac{x+y}{2x^2}$

2) $z = \dfrac{x}{2x^2} + \dfrac{y}{2x^2}$

3) $dz = \dfrac{\partial z}{\partial x}dx + \dfrac{\partial z}{\partial y}dy$

4) $\dfrac{\partial z}{\partial x} = \dfrac{\partial}{\partial x}\left(\dfrac{x}{2x^2} + \dfrac{y}{2x^2}\right) = \dfrac{2x^2 - 4x^2}{(2x^2)^2} + \dfrac{-4xy}{(2x^2)^2}$

$= \dfrac{2x^2 - 4x^2 - 4xy}{4x^4} = \dfrac{x - 2x - 2y}{2x^3}$

5) $\dfrac{\partial z}{\partial y} = \dfrac{\partial}{\partial y}\left(\dfrac{x}{2x^2} + \dfrac{y}{2x^2}\right) = \dfrac{\partial}{\partial y}\left(\dfrac{y}{2x^2}\right) = \dfrac{1}{2x^2}$

6) $dz = \dfrac{-(x+2y)}{2x^3}dx + \dfrac{1}{2x^2}dy$

25

8.3.1 Finding Total Derivatives from Differentials

Given

1) $y = f(x_1, x_2, \ldots, x_n)$

Total differential dy is equal to the sum of the partial changes in y :

2) $dy = \dfrac{\partial y}{\partial x_1}dx_1 + \dfrac{\partial y}{\partial x_2}dx_2 + \ldots + \dfrac{\partial y}{\partial x_n}dx_n$

3) $dy = f_1 dx_1 + f_2 dx_2 + \ldots + f_n dx_n$

The partial total derivative of y wrt x_1, for example, is found by dividing both sides by dx_1

4) $\dfrac{dy}{dx_1} = f_1 + f_2\dfrac{dx_2}{dx_1} + \ldots + f_n\dfrac{dx_n}{dx_1}$

26

8.4 Multivariate Taylor Series

- Recall Taylor's series formula

$$f(x) \approx T(x,c) = f(c) + \frac{f^{/}(c)}{1!}(x-c)^1 + \frac{f^{//}(c)}{2!}(x-c)^2 + \ldots + \frac{f^{(n)}(c)}{n!}(x-c)^n$$

- We want to generalize the Taylor polynomial to multivariate functions. A similar logic to the univariate case gives us:

$$f(x) \approx T(x,a) = f(\mathrm{a}) + Df(\mathrm{a})\,(\mathrm{x}-\mathrm{a})^1 + \frac{1}{2!}(\mathrm{x}-\mathrm{a})^T H(\mathrm{a})(\mathrm{x}-\mathrm{a}) + \ldots$$

- Using abbreviated notation:

$$\mathrm{T}(\mathbf{x},\mathbf{a}) = \Sigma_{j=0 \text{ to } n}\,(1/j!)\, D^j f(\mathbf{a})\,(\mathbf{x}-\mathbf{a})^j.$$

27

8.4 Multivariate Taylor Series

Example: 1st-order Taylor series, around $\mathbf{a} = (d, c) = (0, 0)$ of

$$f(\mathrm{x}, \mathrm{y}) = [(1+\mathrm{x})/(1+\mathrm{y})] - 1$$

$f(\mathrm{x}, \mathrm{y}) = [(1+\mathrm{x})/(1+\mathrm{y})] - 1 \quad \Rightarrow f(c=0, d=0) = [(1+0)/(1+0)] - 1 = 0$

$f_x = 1/(1+\mathrm{y}) \quad \Rightarrow f_x(c=0, d=0) = 1$

$f_y = (-1)(1+\mathrm{x})/(1+\mathrm{y})^2 \quad \Rightarrow f_y(c=0, d=0) = -1$

Then, 1st-order Taylor series formula:

$$f(\mathrm{x}, \mathrm{y}) \approx \mathrm{T}(\mathrm{x}, \mathrm{y}; \mathbf{0}) = 0 + 1\,(\mathrm{x}-0) + (-1)\,(\mathrm{y}-0) = \mathrm{x} - \mathrm{y}$$

- Application to Relative Purchasing Power Parity (PPP):

$$e_{f,T}^{PPP} = [(1+I_d)/(1+I_f)] - 1 \approx (I_d - I_f),$$

where $e_{f,T}$ is the percentage change in exchange rates from t to T, or:

$$e_{f,T} = (S_{t+T}/S_t) - 1.$$

28

8.5 Homogeneous Functions

- Definition:

A function $f(x_1, ..., x_n)$ is *homogeneous of degree r* if multiplication of each of its independent variables by a constant j will alter the value of the function by the proportion j^r, that is;

if $f(jx_1, ..., jx_n) = j^r f(x_1, ... x_n)$, for all $f(jx_1, ..., jx_n)$ in the domain of f

Special cases:

- If $r = 0$, $j^0 = 1$, the function is homogeneous of degree zero

- If $r = 1$, $j^1 = j$, the function is homogeneous of degree one, sometimes called *linearly homogeneous*.

Note: Technically if $j > 0$, we say *positive homogenous.*

29

8.5 Homogeneous Functions: Examples

Examples:

- In applied work, it is common to see homogenous production functions. For example, a firm increases inputs by k, then output increases by k. Then, $f(.)$ is homogenous of degree ($r = 1$), we say, $f(.)$ shows *constant returns to scale*. If $r > 1$ ($r < 1$), $f(.)$ shows *increasing (decreasing) returns to scale.*

- Demand functions are homogeneous. If all prices and income change by the same amount (the budget constraint does not change), the demands remain unchanged. That is,

$$D(jp_1, ..., jp_n, jI) = D(x_1, ... x_n, I) \quad \Rightarrow \text{homogenous of degree 0.}$$

Since individual demands have $r = 0$, the aggregate demand (sum of individual demands) also has $r = 0$.

30

8.5 Homogeneous Functions: Cobb-Douglas

- A popular production function is the Cobb-Douglas:

$$Q = A\, K^{\alpha} L^{\beta}.$$

The Cobb-Douglas function is homogeneous of degree $\alpha + \beta$:

$$A(jK)^{\alpha}(jL)^{\beta} = Aj^{\alpha} j^{\beta} K^{\alpha} L^{\beta}$$
$$= Aj^{\alpha+\beta} K^{\alpha} L^{\beta} = j^{\alpha+\beta} AK^{\alpha} L^{\beta}$$
$$= j^{\alpha+\beta} Q$$

Cases: $\alpha + \beta > 1$ increasing returns (paid < share)

$\alpha + \beta < 1$ decreasing returns (paid > share)

$\alpha + \beta = 1$, constant returns (function is linearly homogeneous)

Note: In empirical work it is usually found that $\alpha+\beta$ are close to 1. Assuming linear homogeneity is common.

31

8.5 Homogeneous Functions: Cobb-Douglas

- Linear homogeneity of $Q = A\, K^{\alpha} L^{\beta}$

- If $\alpha + \beta = 1$, the Cobb-Douglas function is linearly homogeneous. Let $j = 1/L$, then the average physical product of labor (APP_L) and of capital (APP_K) can be expressed as the capital-labor ratio, $k \equiv K/L$:

$$jQ = \frac{Q}{L} = \phi(k) = A\left(\frac{K}{L}\right)^{\alpha}\left(\frac{K}{L}\right)^{1-\alpha} = Ak^{\alpha}$$

$$APP_L = \frac{Q}{L} = \phi(k) = Ak^{\alpha}$$

$$APP_K = \frac{Q}{L}\frac{L}{K} = \frac{\phi(k)}{k} = Ak^{\alpha-1}$$

Note: This result applies to linearly homogeneous functions

$$Q = f(K, L) \Rightarrow \quad jQ = Q/L = f(K/L, L/L) = f(k,1) = \phi(k)$$

32

8.5 Homogeneous Functions: Properties

- Given a linearly homogeneous production function Q = $f(K, L)$, the marginal physical products MPP_L and MPP_K can be expressed as functions of k alone:

$$MPP_K = \phi'(k)$$
$$MPP_L = \phi(k) - k\phi'(k)$$

33

8.5 Homogeneous Functions: Euler's Theorem

Euler's Theorem
Let $f : R^n_+ \to R$ be continuous, and differentiable on R^n_+. Then, f is homogeneous of degree r if and only if for all $\boldsymbol{x} \in \mathbf{R}^n_+$:

$$r\,f(x_1, ..., x_n) = \Sigma_i\, f_i\, x_i$$

Example: Suppose Q = $f(K, L)$, is homogeneous of degree 1, then,

$$Q = K\frac{\partial Q}{\partial K} + L\frac{\partial Q}{\partial L} = K(MPK) + L(MPL)$$

Then, if each input is paid the amount of its marginal product the total product will be exactly exhausted by the distributive shares for all the inputs –i.e., no residual is left.

34

8.5 Homogeneous Functions: Euler's Theorem

• Euler's theorem has a useful corollary:
Suppose that $f: R^n \rightarrow R$ is differentiable and homogeneous of degree k. Then its first-order partial derivatives f_i are homogeneous of degree $k - 1$.

Example: Suppose Q = f(K, L) is homogeneous of degree 1, then, the MPL and MPK are homogeneous of degree 0. This implies:

$$0 = \frac{\partial Q}{\partial L} = L\frac{\partial Q}{\partial L}(\frac{\partial Q}{\partial L}) + K\frac{\partial Q}{\partial K}(\frac{\partial Q}{\partial L}) = L\frac{\partial^2 Q}{\partial L^2} + K\frac{\partial^2 Q}{\partial K \partial L}$$

$$\frac{\partial^2 Q}{\partial K \partial L} = -\frac{L}{K}\frac{\partial^2 Q}{\partial L^2}$$

which is positive since f_{ii} <0. That is, the marginal productivity of one factor increases when the other factor also increases (*Wicksell's law*).

35

8.5 Homogeneous Functions: Black-Scholes

- Recall the BS call pricing formula:

$$C_t = S_t\, N(d1) - K\, e^{-i(T-t)}N(d2)$$

It is easy to check that it is homogenous of degree one in S_t and K, if we assume the other variables are fixed, especially σ (this assumption *is called "sticky-by-strike"*).

We can apply Euler's Theorem to quickly derive Δ:

$$C_t = S_t * \frac{dC_t}{dS_t} + K * \frac{dC_t}{dK}$$

$$\Rightarrow \qquad \Delta = \frac{dC_t}{dS_t} = N(d1).$$

<u>Note</u>: The homogeneity property (in financial engineering "*sticky moneyness regime*") holds for a more general class of pricing models.

36

8.6 Implicit Function Theorem

- So far, if we were given $F(y, x)=0 \rightarrow y = f(x)$.
  - dy/dx easy to calculate (not always realistic situation.)
- Suppose $F(y, x) = x^3 - 2x^2y + 3xy^2 - 22 = 0$,
  - not easy to solve for $y = f(x)$ $\Rightarrow dy/dx = ?$

- *Implicit Function Theorem*: given $F(y, x_1 \ldots, x_m) = 0$

  a) if F has continuous partial derivatives $F_y, F_1, \ldots, F_m$ and $F_y \neq 0$

  b) if at point $(y_0, x_{10}, \ldots, x_{m0})$, we can construct a neighborhood (N) of $(x_1 \ldots, x_m)$, say, by limiting the range of y, $y = f(x_1 \ldots, x_m)$ --i.e., each vector of x's $\rightarrow$ unique y

  Then, i) y is an implicitly defined function $y = f(x_1 \ldots, x_m)$ and

  ii) still satisfies $F(y, x_1 \ldots x_m)$ for every m-tuple in the N such that $F \equiv 0$.

37

8.6.1 Implicit Function Rule

- If the function $F(y, x_1, x_2, \ldots, x_n) = k$ is an implicit function of $y = f(x_1, x_2, \ldots, x_n)$, then

$$F_y dy + F_{x_1} dx_1 + F_{x_2} dx_2 + \ldots + F_{x_n} d_{x_n} = 0$$

where $F_y = \partial F/\partial y$; $F_{x1} = \partial F/\partial x_1$

- From this result, we derive the *implicit function rule.*
- Total differentiation of $F(y, x_1, x_2 \ldots x_n) = 0$, & set $dx_{2 \text{ to } n} = 0$

$$F_y dy = -F_{x_1} dx_1 - F_{x_2} dx_2 - \ldots - F_{x_n} d_{x_n}$$

$$\frac{dy}{dx_1}\Big|_{dx_x, \cdots dx_n = 0} = \frac{\partial y}{\partial x_1} = -\frac{F_{x_1}}{F_y} \qquad \textit{(Implicit function rule)}$$

38

8.6.1 Implicit function problem - Examples

- Given the equation $F(y, x) = x^3 - 2x^2y + 3xy^2 - 22 = 0$,
- *Q1*: Find the implicit function $y = f(x)$ defined at $(y = 3, x = 1)$
- The function F has continuous partial derivatives $F_y, F_1, \ldots, F_m$:

  $\partial F/\partial y = F_y = -2x^2 + 6xy$ $\quad\quad$ $\partial F/\partial x = F_x = 3x^2 - 4xy + 3y^2$
- At $(y_0, x_{10}, \ldots, x_{m0})$ satisfying $F(y, x_1 \ldots, x_m) = 0$, $F_y \neq 0$:

  $F(y = 3, x = 1) = 1^3 - 2 * 1^2 * 3 + 3 * 1 * 3^2 - 22 = 0$;

  $F_y\ (y = 3, x = 1) = -2x^2 + 6xy = -2 * 1^2 + 6 * 1 * 3 = 16$.
- Yes! We have a continuous function f with continuous partial derivatives.

- *Q2*: Find dy/dx by the implicit-function rule. Evaluate it at $(y=3, x=1)$
- $dy/dx = -F_x/F_y = -(3x^2 - 4xy + 3y^2)/(-2x^2 + 6xy)$

  $dy/dx = -(3*1^2 - 4*1*3 + 3*3^2)/(-2*1^2 + 6*1*3) = -18/16 = -9/8$ 39

8.6.2 Derivatives of implicit functions - Examples

- **Example 1**:

If $F(z, x, y) = x^2z^2 + xy^2 - z^3 + 4yz = 0$,

then, $$\frac{\partial y}{\partial z} = -\frac{F_z}{F_y} = -\frac{2x^2z - 3z^2 + 4y}{2xy + 4z}$$

- **Example 2**: Implicit Production function: F (Q, K, L)

  $\partial F/\partial J = F_J$ $\quad\quad$ J = Q, K, L

Applying the implicit function rule:

$\partial Q/\partial L = -(F_L/F_Q)$ $\quad$ - MPP_L

$\partial Q/\partial K = -(F_K/F_Q)$ $\quad$ - MPP_K

$\partial K/\partial L = -(F_L/F_K)$ $\quad$ - MRTS: Slope of the isoquant 40

8.6.3 Extension: Simultaneous equations case

- We have a set of m implicit equations. We are interested in the effect of the exogenous variables (**x**) on the endogenous variables (**y**). That is, dy_i/dx_j.
- Find total differential of each implicit function.
- Let all the differentials $dx_i = 0$ except dx_1 and divide each term by dx_1 (note: dx_1 is a choice)
- Rewrite the system of partial total derivatives of the implicit functions in matrix notation

41

8.6.3 Extension: Simultaneous equations case

Example : 2x2 System

1) $F_1(y_1, y_2, x_1) \equiv 0$

2) $F_2(y_1, y_2, x_2) \equiv 0$

3) $\frac{\partial F_1}{\partial y_1} dy_1 + \frac{\partial F_1}{\partial y_2} dy_2 + \frac{\partial F_1}{\partial x_1} dx_1 = 0$

4) $\frac{\partial F_2}{\partial y_1} dy_1 + \frac{\partial F_2}{\partial y_2} dy_2 + \frac{\partial F_2}{\partial x_2} dx_2 = 0$

5) $\frac{\partial F_1}{\partial y_1} dy_1 + \frac{\partial F_1}{\partial y_2} dy_2 = -\frac{\partial F_1}{\partial x_1} dx_1 + 0 dx_2$

6) $\frac{\partial F_2}{\partial y_1} dy_1 + \frac{\partial F_2}{\partial y_2} dy_2 = \quad 0 dx_1 \quad -\frac{\partial F_2}{\partial x_2} dx_2$

42

8.6.3 Extension: Simultaneous equations case

- Rewrite the system of partial total derivatives of the implicit functions in matrix notation (**Ax** = **d**)

$$7)\ \frac{\partial F_1}{\partial y_1}\frac{dy_1}{dx_1}+\frac{\partial F_1}{\partial y_2}\frac{dy_2}{dx_1}=-\frac{\partial F_1}{\partial x_1} \qquad 10)\ \frac{\partial F_1}{\partial y_1}\frac{dy_1}{dx_2}+\frac{\partial F_1}{\partial y_2}\frac{dy_2}{dx_2}=0$$

$$8)\ \frac{\partial F_2}{\partial y_1}\frac{dy_1}{dx_1}+\frac{\partial F_2}{\partial y_2}\frac{dy_2}{dx_1}=0 \qquad 11)\ \frac{\partial F_2}{\partial y_1}\frac{dy_1}{dx_2}+\frac{\partial F_2}{\partial y_2}\frac{dy_2}{dx_2}=-\frac{\partial F_2}{\partial x_2}$$

$$9)\begin{bmatrix}\frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2}\\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2}\end{bmatrix}\begin{bmatrix}\frac{dy_1}{dx_1}\\ \frac{dy_2}{dx_1}\end{bmatrix}=\begin{bmatrix}-\frac{\partial F_1}{\partial x_1}\\ 0\end{bmatrix} \qquad 12)\begin{bmatrix}\frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2}\\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2}\end{bmatrix}\begin{bmatrix}\frac{dy_1}{dx_2}\\ \frac{dy_2}{dx_2}\end{bmatrix}=\begin{bmatrix}0\\ -\frac{\partial F_2}{\partial x_2}\end{bmatrix}$$

43

8.6.3 Extension: Simultaneous equations case

- Solve the comparative statics of endogenous variables in terms of exogenous variables using Cramer's rule

$$13)\begin{bmatrix}\frac{dy_1}{dx_1}\\ \frac{dy_2}{dx_1}\end{bmatrix}=\begin{bmatrix}\frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2}\\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2}\end{bmatrix}^{-1}\begin{bmatrix}-\frac{\partial F_1}{\partial x_1}\\ 0\end{bmatrix} \qquad 14)\begin{bmatrix}\frac{dy_1}{dx_2}\\ \frac{dy_2}{dx_2}\end{bmatrix}=\begin{bmatrix}\frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2}\\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2}\end{bmatrix}^{-1}\begin{bmatrix}0\\ -\frac{\partial F_2}{\partial x_2}\end{bmatrix}$$

$$\frac{dy_1}{dx_1}=\frac{\begin{vmatrix}-\frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial y_2}\\ 0 & \frac{\partial F_2}{\partial y_2}\end{vmatrix}}{|J|}=\frac{-\frac{\partial F_1}{\partial x_1}\text{ x }\frac{\partial F_2}{\partial y_2}}{\frac{\partial F_1}{\partial y_1}\text{ x }\frac{\partial F_2}{\partial y_2}-\frac{\partial F_1}{\partial y_2}\text{ x }\frac{\partial F_2}{\partial y_1}}$$

44

8.7 Application: The Market Model

- Assume the demand and supply functions for a commodity are general form explicit functions
  $Q_d = D(P, Y_0)$ ($D_p < 0$; $D_{Y0} > 0$)
  $Q_s = S(P, T_0)$ ($S_p > 0$; $S_{T0} < 0$)
- Q is quantity, P is price, (endogenous variables)
  Y_0 is income, T_0 is the tax (exogenous variables)
  no parameters, all derivatives are continuous
- Find $\partial P/\partial Y_0$, $\partial P/\partial T_0$
  $\partial Q/\partial Y_0$, $\partial Q/\partial T_0$
- Solution:
  - Either take total differential or apply implicit function rule
  - Use the partial derivatives as parameters
  - Set up structural form equations as **A**x = d,
  - Invert A matrix or use Cramer's rule to solve for $\partial x/\partial d$

45

8.7 Application: The Market Model

1) $D(\overline{P}, Y_0) = S(\overline{P}, T_0) \equiv \overline{Q}$

2) $F^1(P, Q; Y_0, T_0) = D(\overline{P}, Y_0) - \overline{Q} \equiv 0$

3) $F^2(P, Q; Y_0, T_0) = S(\overline{P}, T_0) - \overline{Q} \equiv 0$

Suppose we are interested in finding $d\overline{Q}/dY_0$.

Take the total differential of equations (2) & (3) and organize;

4) $D_P^{/} d\overline{P} - d\overline{Q} = -D_{Y_0}^{/} dY_0$

5) $S_P^{/} d\overline{P} - d\overline{Q} = -S_{T_0}^{/} dT_0$

Put equations (4) & (5) in matrix format (Ax = d);

6) $\begin{bmatrix} D_P^{/} & -1 \\ S_P^{/} & -1 \end{bmatrix} \begin{bmatrix} d\overline{P} \\ d\overline{Q} \end{bmatrix} = \begin{bmatrix} -D_{Y_0}^{/} & 0 \\ 0 & -S_{T0}^{/} \end{bmatrix} \begin{bmatrix} dY_0 \\ dT_0 \end{bmatrix} = \begin{bmatrix} -D_{Y_0}^{/} dY_0 \\ -S_{T0}^{/} dT_0 \end{bmatrix}$

46

8.7 Application: The Market Model

6) $$\begin{bmatrix} D_P' & -1 \\ S_P' & -1 \end{bmatrix}\begin{bmatrix} d\overline{P} \\ d\overline{Q} \end{bmatrix} = \begin{bmatrix} -D_{Y_0}' dY_0 \\ -S_{T0}' dT_0 \end{bmatrix}$$

Take the partial total derivative of equation (6) wrt to dY_0.

7) $$\begin{bmatrix} \dfrac{d\overline{P}}{dY_0} \\ \dfrac{d\overline{Q}}{dY_0} \end{bmatrix} = \begin{bmatrix} D_P' & -1 \\ S_P' & -1 \end{bmatrix}^{-1} \begin{bmatrix} -D_{Y_0}' \\ 0 \end{bmatrix}$$

We want to calculate : $\dfrac{d\overline{Q}}{dY_0} = \dfrac{|J_2|}{|J|}$

Calculate the Jacobian determinant, $|J|$, and $|J_2|$.

$$|J| = \begin{vmatrix} D_P' & -1 \\ S_P' & -1 \end{vmatrix} = S_P' - D_P' > 0 \quad ; \qquad |J_2| = \begin{vmatrix} D_P' & -D_{Y_0}' \\ S_P' & 0 \end{vmatrix} = S_P' D_{Y_0}' > 0$$

$$\frac{d\overline{Q}}{dY_0} = \frac{S_P' D_{Y_0}'}{S_P' - D_P'} > 0.$$

47

8.8 Limitations of Comparative Statics

- Comparative statics answers the question: how does the equilibrium change with a change in a parameter.
- The adjustment process is ignored
- New equilibrium may be unstable
- Before dynamic, optimization

48

8.9 Cheat-Sheet: Rules for Vector Derivatives

- Consider the linear function: $y = f(\boldsymbol{x}) = \boldsymbol{x}'\, \beta + \omega$

where $\boldsymbol{x}$ and β are k-dimensional vectors and ω is a constant.

We derive the gradient in matrix notation as follows:

1. Convert to summation notation: $f(\boldsymbol{x}) = \sum_i^k x_i\, \beta_i$

2. Take partial derivative w.r.t. element x_j: $\frac{\partial}{\partial x_j}\left[\sum_i^k x_i\, \beta_i\right] = \beta_j$

3. Put all the partial derivatives in a vector:

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\boldsymbol{x})}{\partial x_k} \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}$$

4. Convert to matrix notation: $\nabla f(\boldsymbol{x}) = \beta$

49

8.9 Cheat-Sheet: Rules for Vector Derivatives

- Consider a quadratic form: $q = f(\boldsymbol{x}) = \boldsymbol{x}'\, \mathbf{A}\, \boldsymbol{x}$

where $\boldsymbol{x}$ is kx1 vector and $\mathbf{A}$ is a kxk matrix, with a_{ji} elements.

Steps:

1. Convert to summation notation:

$$f(\boldsymbol{x}) = \boldsymbol{x}' \begin{bmatrix} \sum_i^k a_{j1} x_j \\ \vdots \\ \sum_i^k a_{jk} x_j \end{bmatrix} = \sum_i^k \sum_i^k x_i\, a_{ji}\, x_j$$

(we rewrite $\sum_i^k \sum_i^k x_i\, a_{ji} x_j = \sum_i^k a_{ii} x_i^2 + \sum_i^k \sum_{i \neq j}^k x_i\, a_{ji} x_j$)

2. Take partial derivative w.r.t. element x_j:

$$\frac{\partial}{\partial x_j}\left[\sum_i^k \sum_i^k x_i\, a_{ji} x_j\right] = 2\, a_{jj}\, x_j + \sum_{i \neq k}^k x_i\, a_{ij} + \sum_{i \neq k}^k a_{ji}\, x_j$$

50

8.9 Cheat-Sheet: Rules for Vector Derivatives

2. Take partial derivative w.r.t. element x_j:

$$\frac{\partial}{\partial x_j}\left[\sum_i^k \sum_i^k x_i\, a_{ji} x_j\right] = 2\, a_{jj}\, x_j + \sum_{i\neq k}^k x_i\, a_{ij} + \sum_{i\neq k}^k a_{ji}\, x_i$$

$$= \sum_i^k x_i\, a_{ij} + \sum_i^k a_{ji}\, x_i$$

3. Put all the partial derivatives in a vector:

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\boldsymbol{x})}{\partial x_k} \end{bmatrix} = \begin{bmatrix} \sum_i^k x_i\, a_{i1} \\ \vdots \\ \sum_i^k x_i\, a_{ik} \end{bmatrix} + \begin{bmatrix} \sum_i^k a_{1i}\, x_i \\ \vdots \\ \sum_i^k a_{ki}\, x_i \end{bmatrix}$$

4. Convert to matrix notation:

$$\nabla f(\boldsymbol{x}) = \mathbf{A}'\,\boldsymbol{x} + \mathbf{A}\,\boldsymbol{x} = (\mathbf{A}' + \mathbf{A})\,\boldsymbol{x}$$

If **A** is symmetric, then $\nabla f(\boldsymbol{x}) = 2\,\mathbf{A}\,\boldsymbol{x}$

51

8.9 Cheat-Sheet: Rules for Vector Derivatives

- Hessian of a linear function and a quadratic form
- Linear function: $\mathbf{y} = f(\boldsymbol{x}) = \boldsymbol{x}'\,\beta + \omega$

We have already derived: $\nabla f(\boldsymbol{x}) = \beta$

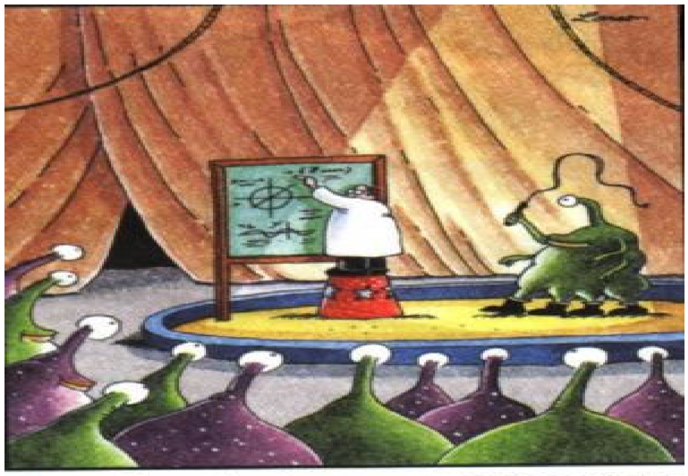
Then, $\mathbf{H} = \frac{\partial}{\partial \boldsymbol{x}}[\nabla f(\boldsymbol{x}) = \beta] = \mathbf{0}$

- Quadratic form: $\mathbf{q} = f(\boldsymbol{x}) = \boldsymbol{x}'\,\mathbf{A}\,\boldsymbol{x}$

We have already derived $\nabla f(\boldsymbol{x}) = (\mathbf{A}' + \mathbf{A})\,\boldsymbol{x}$

Then, $\mathbf{H} = \frac{\partial}{\partial \boldsymbol{x}}[\nabla f(\boldsymbol{x}) = (\mathbf{A}' + \mathbf{A})\,\boldsymbol{x}\;] = (\mathbf{A}' + \mathbf{A})$

52



Abducted by an alien circus company, Professor Doyle is forced to write calculus equations in center ring.