Mathematics for Economists: Lecture 4

Juuso Välimäki

Aalto University School of Business

Spring 2022

Content of Lecture 4

- In Lecture 3, Linear approximations of non-linear functions
- This Lecture:
  1. Linear approximation of vector valued functions
  2. Implicit function theorem
  3. Comparative statics of economic models

Linear approximation of vector-valued functions

- What is a vector valued function?
  - A function whose values take the form of a column vector
  - Each component in the vector is a (possibly) multivariate function

$$\boldsymbol{f} = \begin{pmatrix} f_1(x_1, ..., x_n) \\ f_2(x_1, ..., x_n) \\ \vdots \\ f_n(x_1, ..., x_n) \end{pmatrix}.$$

- What is an economic example: the vector of demands

$$\boldsymbol{x}(p_1, ..., p_n, I) = \begin{pmatrix} x_1(p_1, ..., p_n, I) \\ x_2(p_1, ..., p_n, I) \\ \vdots \\ x_n(p_1, ..., p_n, I) \end{pmatrix}.$$

- The domain of this function is $\{(p_1, ..., p_n, I)|p_i > 0 \text{ for all } i, \text{ and } I > 0\}$.
- The values of this function are in $\{(x_1, ..., x_n)|x_i \geq 0 \text{ for all } i\}$.

Linear approximation of vector-valued functions

- How do we find a linear approximation?
  - Vector of linear approximations to component functions

$$D_{\boldsymbol{x}}\boldsymbol{f} = \begin{pmatrix} D_{\boldsymbol{x}} f_1(x_1, \ldots, x_n) \\ D_{\boldsymbol{x}} f_2(x_1, \ldots, x_n) \\ \vdots \\ D_{\boldsymbol{x}} f_n(x_1, \ldots, x_n) \end{pmatrix}.$$

Linear approximation of vector-valued functions

- Writing in full:

$$
D_{\boldsymbol{x}} \boldsymbol{f} = \begin{pmatrix} \frac{\partial f_1(x_1,\ldots,x_n)}{\partial x_1} & \frac{\partial f_1(x_1,\ldots,x_n)}{\partial x_2} & \cdots & \frac{\partial f_1(x_1,\ldots,x_n)}{\partial x_n} \\ \frac{\partial f_2(x_1,\ldots,x_n)}{\partial x_1} & \frac{\partial f_2(x_1,\ldots,x_n)}{\partial x_2} & \cdots & \frac{\partial f_2(x_1,\ldots,x_n)}{\partial x_n}) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_n(x_1,\ldots,x_n)}{\partial x_1} & \frac{\partial f_n(x_1,\ldots,x_n)}{\partial x_2} & \cdots & \frac{\partial f_n(x_1,\ldots,x_n)}{\partial x_n} \end{pmatrix}.
$$

- We get the linear approximation at $\hat{\boldsymbol{x}}$ by evaluating the derivative matrix at $\hat{\boldsymbol{x}}$:

$$
D_{\boldsymbol{x}} \boldsymbol{f}(\hat{\boldsymbol{x}}) = \begin{pmatrix} D_{\boldsymbol{x}} f_1(\hat{\boldsymbol{x}}) \\ D_{\boldsymbol{x}} f_2(\hat{\boldsymbol{x}}) \\ \vdots \\ D_{\boldsymbol{x}} f_n(\hat{\boldsymbol{x}}) \end{pmatrix}.
$$

Linear approximation of vector-valued functions: numerical example

- Consider the following vector-valued function:

$$\boldsymbol{f}(x,y,z) = \begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \end{pmatrix} = \begin{pmatrix} x + y^2 + \frac{1}{z} \\ -x + \sqrt{y} + 2z \end{pmatrix}.$$

- To compute the derivative at $(x = 1, y = 1, z = 1)$, compute first the matrix of partial derivatives:

$$D_{x,y,z}\boldsymbol{f}(x,y,z) = \begin{pmatrix} 1 & +2y & -\frac{1}{z^2} \\ -1 & \frac{1}{2\sqrt{y}} & 2 \end{pmatrix}.$$

- Evaluating at $(1,1,1)$ gives

$$D_{x,y,z}\boldsymbol{f}(1,1,1) = \begin{pmatrix} 1 & 2 & -1 \\ -1 & \frac{1}{2} & 2 \end{pmatrix}.$$

Chain rule for multivariate functions:

- Recall the chain rul: If $y = f(x)$ and $z = g(y)$, then for $h(x) = g(f(x))$:

$$h'(x) = g'(f(x))f'(x).$$

- Consider now a similar situation where $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^k$, and $\boldsymbol{g} : \mathbb{R}^k \to \mathbb{R}^m$ and $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$ given by $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{f}(\boldsymbol{x}))$. Let $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$.
- By the linear approximation at $\hat{\boldsymbol{x}}$ to $\boldsymbol{f}$, we get:

$$\boldsymbol{f}(\hat{\boldsymbol{x}} + \Delta\boldsymbol{x}) \approx \boldsymbol{f}(\hat{\boldsymbol{x}}) + D_{\boldsymbol{x}}\boldsymbol{f}(\hat{\boldsymbol{x}})\Delta\boldsymbol{x}.$$

- Similarly

$$\boldsymbol{g}(\boldsymbol{f}(\hat{\boldsymbol{x}}) + D_{\boldsymbol{x}}(\hat{\boldsymbol{x}})\Delta\boldsymbol{x}) \approx \boldsymbol{g}(\boldsymbol{f}(\hat{\boldsymbol{x}})) + D_{\boldsymbol{y}}\boldsymbol{g}(f(\hat{\boldsymbol{x}}))D_{\boldsymbol{x}}\boldsymbol{f}(\hat{\boldsymbol{x}})\Delta\boldsymbol{x}.$$

- Hence we have:

$$D_{\boldsymbol{x}}\boldsymbol{h}(\hat{\boldsymbol{x}}) = D_{\boldsymbol{y}}\boldsymbol{g}(\boldsymbol{f}(\hat{\boldsymbol{x}}))D_{\boldsymbol{x}}\boldsymbol{f}(\hat{\boldsymbol{x}})\Delta\boldsymbol{x}.$$

Comparative statics: motivating examples

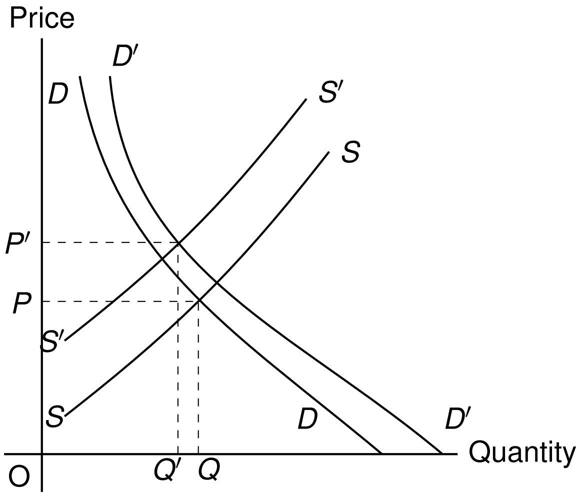

Figure: Exogenous variable shifting demand and supply

Comparative statics: motivating examples

- In Principles 1, we argued that at optimal consumption,

$$MRS_{x_1,x_2}(\hat{\boldsymbol{x}}) = \frac{p_1}{p_2},$$

  where p_i is the price of good i.
- We have also the budget constraint:

$$p_1 x_1 + p_2 x_2 = w,$$

  where w is the total budget.

$$\begin{aligned} p_2 \tfrac{\partial u(x_1,x_2)}{\partial x_1} - p_1 \tfrac{\partial u(x_1,x_2)}{\partial x_2} &= 0, \\ p_1 x_1 + p_2 x_2 - w &= 0. \end{aligned}$$

- Again for many u, no explicit solution is possible.
- Still, how do the optimal consumptions change when some of the p_1, p_2, w change?

Linear implicit function theorem

- Because of linearity, this is not really needed since the system can be solved explicitly
- Consider the system of equations:

$$\begin{aligned} a_{11}y_1 + ... + a_{1n}y_n + b_{11}x_1 + ... + b_{1m}x_m &= 0, \\ &\vdots \\ a_{n1}y_1 + ... + a_{nn}y_n + b_{n1}x_1 + ... + b_{nm}x_m &= 0. \end{aligned}$$

- In matrix form:

$$\boldsymbol{f}(\boldsymbol{y};\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{B}\boldsymbol{x} = 0,$$

where $\boldsymbol{A}$ is an $n \times n$ matrix and $\boldsymbol{B}$ is an $n \times m$ matrix, $\boldsymbol{y} = (y_1, ..., y_n)$, $\boldsymbol{x} = (x_1, ..., x_m)$.

Linear implicit function theorem

- Assume that the system is solved at $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{x}})$:

$$\boldsymbol{f}\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right) = 0 \text{ or } \boldsymbol{A}\widehat{\boldsymbol{y}} + \boldsymbol{B}\widehat{\boldsymbol{x}} = 0,$$

  and consider the effect of a small change $(d\boldsymbol{y}; d\boldsymbol{x}) = (dy_1, ..., dy_n; dx_1, ..., dx_m)$ on the value of f :

$$\boldsymbol{f}\left(\widehat{\boldsymbol{y}} + d\boldsymbol{y}, \widehat{\boldsymbol{x}} + d\boldsymbol{x}\right) - \boldsymbol{f}\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right) = \boldsymbol{A}d\boldsymbol{y} + \boldsymbol{B}d\boldsymbol{x}$$

$$= D_{\boldsymbol{y}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right) d\boldsymbol{y} + D_{\boldsymbol{x}}f\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right) d\boldsymbol{x},$$

  where $D_{\boldsymbol{y}}\boldsymbol{f}(\hat{\boldsymbol{y}}, \hat{\boldsymbol{x}})$ consists of the partial derivatives of $\boldsymbol{f}$ w.r.t. the endogenous variables $\boldsymbol{y}$ and $D_{\boldsymbol{x}}\boldsymbol{f}(\hat{\boldsymbol{y}}, \hat{\boldsymbol{x}})$ w.r.t. the exogenous variables x.

Linear implicit function theorem

- For
$$\boldsymbol{f}(\boldsymbol{y};\boldsymbol{x}) = 0.$$
to hold at $(\boldsymbol{y},\boldsymbol{x}) = \left(\widehat{\boldsymbol{y}} + d\boldsymbol{y}, \widehat{\boldsymbol{x}} + d\boldsymbol{x}\right)$, the change must be zero:
$$D_{\boldsymbol{y}}\boldsymbol{f}(\widehat{\boldsymbol{y}};\widehat{\boldsymbol{x}})d\boldsymbol{y} + D_{\boldsymbol{x}}\boldsymbol{f}(\widehat{\boldsymbol{y}};\widehat{\boldsymbol{x}})d\boldsymbol{x} = 0.$$
- In other words,
$$d\boldsymbol{y} = -D_{\boldsymbol{y}}\boldsymbol{f}(\widehat{\boldsymbol{y}};\widehat{\boldsymbol{x}})^{-1}D_{\boldsymbol{x}}\boldsymbol{f}(\widehat{\boldsymbol{y}};\widehat{\boldsymbol{x}})d\boldsymbol{x} = -\boldsymbol{A}^{-1}\boldsymbol{B}d\boldsymbol{x}.$$
- This equation has a solution for all $d\boldsymbol{x}$ if and only if $\boldsymbol{A}^{-1}$ exists, i.e. if $\boldsymbol{A} = D_{\boldsymbol{y}}\boldsymbol{f}(\widehat{\boldsymbol{y}};\widehat{\boldsymbol{x}})$ has full rank.
- The generalization of this result for the non-linear case in a neighborhood of $(\widehat{\boldsymbol{y}};\widehat{\boldsymbol{x}})$ is called the *implicit function theorem*.

Linear implicit function theorem: Example

Consider the linear system around $(\hat{y}_1, \hat{y}_2, \hat{x}) = (2, 5, -3)$:

$$\begin{aligned} 2y_1 + y_2 + 3x &= 0, \\ y_1 - y_2 - x &= 0. \end{aligned}$$

Compute the value of the function at $(\hat{y}_1 + dy_1, \hat{y}_2 + dy, \hat{x} + dx)$:

$$\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 + dy_1 \\ 5 + dy_2 \end{pmatrix} + \begin{pmatrix} 3 \\ -1 \end{pmatrix} (3 + dx) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

or

$$\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} dy_1 \\ dy_2 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix} dx.$$

By Cramer's rule:

$$dy_1 = \frac{\det \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} dx}{\det \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}} = \frac{-2}{3} dx, \quad dy_2 = \frac{\det \begin{pmatrix} 2 & -3 \\ 1 & 1 \end{pmatrix} dx}{\det \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}} = \frac{-5}{3} dx.$$

Implicit function for y: $f(x, y) = x^2 + y^2 = 1$

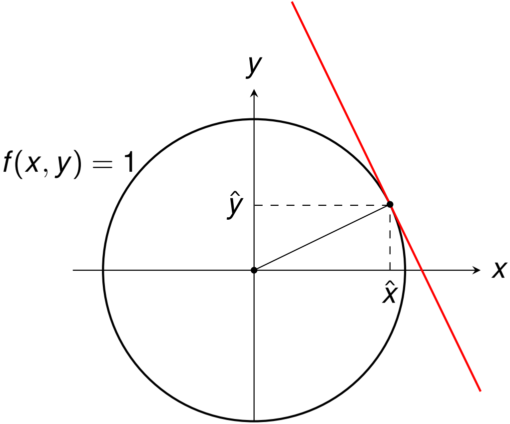

Figure: What's the slope of the tangent at $(\hat{x}, \hat{y})$?

Implicit function theorem for $n = m = 1$

We start this section with an example of a univariate function.

$$f(y, x) = xy + \ln(xy + x) = 0. \tag{1}$$

- Note that $(\widehat{y}, \widehat{x}) = (0, 1)$ satisfies equation 2.
- What is the impact of a small change dx in $\hat{x}$ on the value of y satisfying the equation.
- We are interested in all points (y, x) near $(0, 1)$ satisfying equation 2.
- Let's assume that such a $y(x)$ exists for all x near $\hat{x}$.
- Assume also that $y(x)$ has a derivative at $\hat{x}$. We can then write:

$$g(x) = f(y(x), x) = xy(x) + \ln(xy(x) + x) = 0$$

for all x near $\hat{x} = 1$.

- We see that the original equation has been reduced to an equation in a single variable x.
- Since the composite function is constant in x (=0), the composite function g must have a zero derivative in x near $\hat{x} = 1$.
- By the chain rule:

$$g'(x) = \frac{\partial f(y;x)}{\partial y} y'(x) + \frac{\partial f(y;x)}{\partial x}$$

$$= \left(x + \frac{x}{xy + x}\right) y'(x) + y + \frac{y+1}{xy+x}.$$

- By requiring $g'(1) = 0$, we get:

$$y'(1) = -\frac{\frac{\partial f(0,1)}{\partial x}}{\frac{\partial f(0,1)}{\partial y}} = -\frac{1}{2}.$$

- Notice that this is a valid computation only if $\frac{\partial f(0,1)}{\partial y} \neq 0$.

One-dimensional implicit function theorem

Theorem

Let $f(y,x)$ be a continuously differentiable in a neighborhood of $(\hat{y},\hat{x})$ and $f(\hat{y},\hat{x}) = 0$. If $\frac{\partial f(\hat{y},\hat{x})}{\partial y} \neq 0$, then there exists a continuously differentiable function $y(x)$ in a neighborhood $B_{\hat{x}}$ of $\hat{x}$ such that:

1. *$f(y(x),x) = 0$ for all $x \in B_{\hat{x}}$,*
2. *$y(\hat{x}) = \hat{y}$,*
3. *The derivative of y at $\hat{x}$ satisfies:*

$$y'(\hat{x}) = -\frac{\frac{\partial f(\hat{y},\hat{x})}{\partial x}}{\frac{\partial f(\hat{y},\hat{x})}{\partial y}}$$

The textbook has a proof of this theorem.

The Implicit function theorem

- Consider now a continuously differentiable non-linear function

$$\boldsymbol{f} : \mathbb{R}^{n+m} \to \mathbb{R}^{n}$$

in a neighborhood of the point $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{x}}) \in \mathbb{R}^{n+m}$, where

$$\boldsymbol{f}\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right) = 0.$$

- Use the derivative of $D\boldsymbol{f}(\hat{\boldsymbol{x}}, \hat{\boldsymbol{x}})$ to approximate $\boldsymbol{f}$ at $(\hat{\boldsymbol{y}} + d\boldsymbol{y}, \hat{\boldsymbol{x}} + d\boldsymbol{x})$:

$$\boldsymbol{f}\left(\widehat{\boldsymbol{y}} + d\boldsymbol{y}, \widehat{\boldsymbol{x}} + d\boldsymbol{x}\right) - f\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right) = D\boldsymbol{f}\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right)) d\boldsymbol{x}, d\boldsymbol{y}) = D_{\boldsymbol{y}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right) d\boldsymbol{y} + D_{\boldsymbol{x}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right) d\boldsymbol{x},$$

- Suppose we have a solution to the system at $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{x}})$:

$$D_{\boldsymbol{y}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right) d\boldsymbol{y} + D_{\boldsymbol{x}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right) d\boldsymbol{x} = 0.$$

- Since $D_{\boldsymbol{y}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right)$ ja $D_{\boldsymbol{x}}f\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right)$ are matrices, we continue here exactly as in the linear case.
- With differential calculus, we have reduced the really complicated non-linear problem to the much simpler linear case locally, i.e. in a neighborhood of the solution point $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{x}})$.

The Implicit function theorem: An example

- Consider the following system:

$$\begin{aligned} f_1(y_1, y_2; x_1, x_2) &= y_1 y_2^2 - x_1 x_2 + x_2 + 1 = 0, \\ f_2(y_1, y_2; x_1, x_2) &= y_1 + \frac{x_1}{y_2} + x_2 - 5 = 0. \end{aligned}$$

- Analyze the system of equations in a neighborhood of the point

$$\left(\widehat{y}_1, \widehat{y}_2; \widehat{x}_1, \widehat{x}_2\right) = (1, 1, 2, 2).$$

The Implicit function theorem: An example

- Check first that the equation is satisfied at $(1, 1, 2, 2)$ and form the appropriate matrices of partial derivatives:

$$D_{\mathbf{y}}\mathbf{f}\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right) = \begin{pmatrix} \frac{\partial f_1\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)}{\partial y_1} & \frac{\partial f_1\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)}{\partial y_2} \\ \frac{\partial f_2\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)}{\partial y_1} & \frac{\partial f_2\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)}{\partial y_2} \end{pmatrix} = \begin{pmatrix} \widehat{y}_2^2 & 2\widehat{y}_1\widehat{y}_2 \\ 1 & \frac{-\widehat{x}_1}{\widehat{y}_2^2} \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix},$$

$$D_{\mathbf{x}}\mathbf{f}\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right) = \begin{pmatrix} \frac{\partial f_1\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)}{\partial x_1} & \frac{\partial f_1\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)}{\partial x_2} \\ \frac{\partial f_2\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)}{\partial x_1} & \frac{\partial f_2\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -\widehat{x}_2 & 1-\widehat{x}_1 \\ \frac{1}{\widehat{y}_2} & 1 \end{pmatrix} = \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix}.$$

The Implicit function theorem: An example

- We see that $\det\left(D_{\mathbf{y}}\boldsymbol{f}\left(\widehat{\mathbf{y}};\widehat{\mathbf{x}}\right)\right) \neq 0$, and therefore the matrix $D_{\mathbf{y}}\boldsymbol{f}(\hat{\mathbf{y}},\hat{\mathbf{x}})$ has full rank and an inverse matrix $[D_{\mathbf{y}}\boldsymbol{f}(\hat{\mathbf{y}},\hat{\mathbf{x}})]^{-1}$
- Exercise: Show that

$$[D_{\mathbf{y}}\boldsymbol{f}(\hat{\mathbf{y}},\hat{\mathbf{x}})]^{-1} = \frac{-1}{4}\begin{pmatrix} -2 & -2 \\ -1 & 1 \end{pmatrix},$$

and therefore:

$$d\mathbf{y} = \frac{1}{4}\begin{pmatrix} -2 & -2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix} d\mathbf{x}.$$

Implicit function theorem

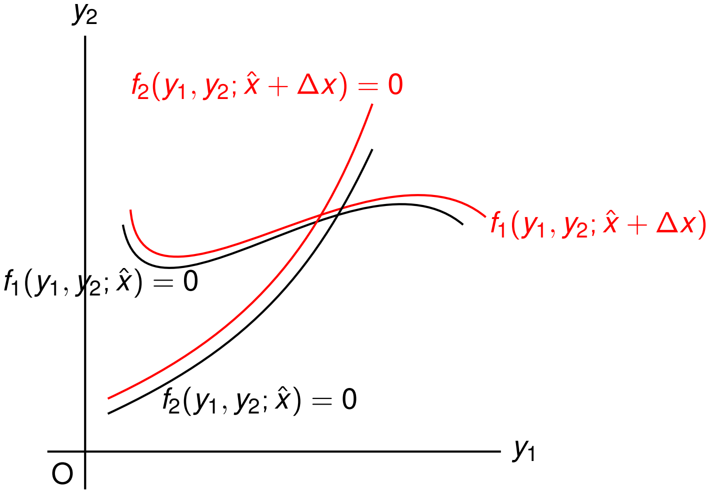

Figure: Implicit function theorem: exogenous changes in x. Red curves after change.

Failure of implicit function theorem

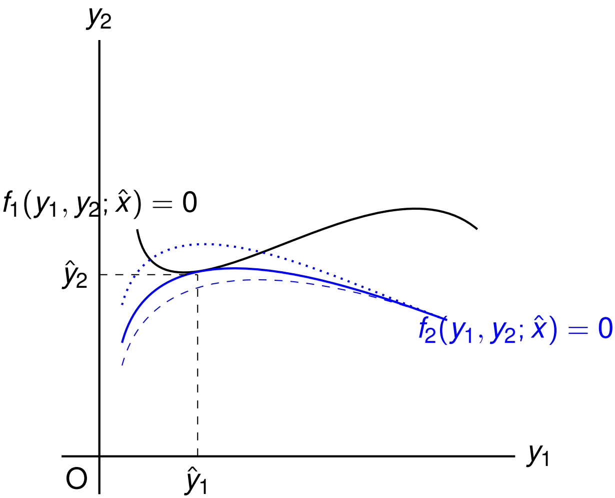

Figure: $D_{\boldsymbol{y}}\boldsymbol{f}(\hat{y}_1, \hat{y}_2; \hat{x})$ does not have full rank. Solid curves drawn for $\hat{x}$.

The Implicit function theorem: Main theorem

We are now ready for the main theorem in this section.

Theorem

Let $\boldsymbol{f} : \mathbb{R}^{n+m} \to \mathbb{R}^n$ *be continuously differentiable in a neighborhood* $B^{\varepsilon}\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right)$, *of* $(\hat{\boldsymbol{y}}, \hat{\boldsymbol{x}})$ *and*

$$\boldsymbol{f}\left(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{x}}\right) = 0.$$

If $\det\left(D_{\boldsymbol{y}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right)\right) \neq 0$, *then there exists a continuously differentiable function* $\boldsymbol{y}(\boldsymbol{x})$ *defined in a neighborhood* $B^{\delta}\left(\widehat{\boldsymbol{x}}\right)$ *of* $\hat{\boldsymbol{x}}$ *such that :*

1. $\boldsymbol{f}(\boldsymbol{y}(\boldsymbol{x}), \boldsymbol{x}) = 0$ *for all* $\boldsymbol{x} \in B^{\delta}\left(\widehat{\boldsymbol{x}}\right)$,
2. $\boldsymbol{y}\left(\widehat{\boldsymbol{x}}\right) = \widehat{\boldsymbol{y}}$,
3. *The derivative of the function* $\boldsymbol{y}$ *satisfies:*

$$D\boldsymbol{y}\left(\widehat{\boldsymbol{x}}\right) = -\left(D_{\boldsymbol{y}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right)\right)^{-1} D_{\boldsymbol{x}}\boldsymbol{f}\left(\widehat{\boldsymbol{y}}; \widehat{\boldsymbol{x}}\right)$$

The Implicit function theorem: Main theorem

- Proving this theorem is beyond the scope of this course.
- Assuming points 1. and 2. above, point 3. is an application of the chain rule in the vector-valued multivariate case.
- It is nothing more than a local version of the linear implicit function theorem.
- Parts 1. and 2. require some more sophisticated mathematics. Proving the existence of the implicit function $y(x)$ near $\hat{x}$ requires the use of a fixed point theorem (similar to the case of showing the existence of local solutions to differential equations).
- We will see more examples once we have more tools from optimization available.

Next lecture:

- Higher order Taylor approximations
- Quadratic forms
- Minima and maxima of non-linear functions
- Applications: Least squares estimators, Cost minimization