Mathematics for Economists: Lecture 4

Juuso Välimäki

Aalto University School of Business

Spring 2020

This lecture covers

1. Utility and production functions
    1.1 Interpretation of partial derivatives
    1.2 MRS and MRTS
2. Motivating example for implicit function theorem
3. Linear implicit function theorem
4. Implicit function theorem for single endogenous variable
5. Full implicit function theorem

Utility functions: underlying preferences

- Motivation: we start with a preference relation on the set of possible consumption vectors
- Possible consumptions $x = (x_1, ..., x_n)$ such that $x_i \geq 0$ for all i. We write $x \in \mathbb{R}^n_+$
- We write $x \succsim y$ if x is at least as good as y for our consumer and $x \succ y$ if x is strictly better
- A preference is rational if for all $x, y \in \mathbb{R}^n_+$ i) Either $x \succsim y$ or $y \succsim x$ or both (i.e. the consumer can make comparisons), ii) If $x \succsim y$ and $y \succsim z$, then $x \succsim z$ (we say that preference is transitive).
- (technical) A preference relation is continuous if $x \succ y$ implies that all consumption vectors close enough to x are strictly better than y (For all y, the set of x such that $x \succ y$ is open).

Utility functions: representing preferences

- A utility function $u : \mathbb{R}^n \to \mathbb{R}$ is said to *represent preferences* $\succsim$ if

$$u(x) \geq u(y) \Leftrightarrow x \succsim y.$$

- In later studies you may see a proof of this important theorem by Debreu

Theorem

If $\succsim$ is a continuous rational preference relation on $\mathbb{R}^n_+$, then there exists a continuous utility function $u : \mathbb{R}^n \to \mathbb{R}$ representing $\succsim$.

- This is the meaning for utility functions in economics.
- Maximizing utility function is equivalent to picking the best alternative according to the underlying preferences

Utility functions: Examples

- Consumption of food x_1 and housing x_2
- Assume that more is strictly better than less for each good (if you have trouble with this assumption, consider x_i to be measured in quality adjusted units)
- What is the mathematical assumption corresponding to this?
- Strictly positive partial derivatives $\frac{\partial u(\hat{x})}{\partial x_i}$ for all $\hat{x}$ and $i \in \{1, 2\}$. These are called the marginal utilities and denoted by $MU_{x_i}(\hat{x})$.
- For example $u(x_1, x_2) = x_1^{\frac{2}{3}} x_2^{\frac{1}{3}}$.
- Recall from Principles 1 that *marginal rate of substitution between* x_1 *and* x_2 was defined the largest amount of y that the consumer would be willing to give to get one extra small unit of x_1.
- How to express mathematically?

Utility functions: Examples

- Consider the directional derivative and level curves and calculate the changes in the utility from point $\hat{x} = (\hat{x}_1, \hat{x}_2)$ to $(\hat{x}_1 + dx_1, \hat{x}_2 + dx_2)$. using the linear approximation, we get

$$u(\hat{x}_1 + dx_1, \hat{x}_2 + dx_2) - u(\hat{x}_1, \hat{x}_2) = Du(\hat{x}_1, \hat{x}_2)(dx_1, dx_2)$$

$$= \left(\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}, \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}\right) \cdot (dx_1, dx_2)$$

$$= \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1} dx_1 + \frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2} dx_2.$$

- This change is therefore zero only if

$$\frac{dx_2}{dx_1} = -\frac{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1}}{\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2}} = -\frac{MU_{x_1}(\hat{x})}{MU_{x_2}(\hat{x})} =: -MRS_{x_1, x_2}(\hat{x}).$$

- Notice that for this to make sense, we must have nonzero marginal utility for x_2, but we already assumed this

Utility functions: Examples

- Compute MRS_{x_1,x_2} at $\hat{x} = (10, 5)$ for $u(x_1, x_2) = x_1^{\frac{2}{3}} x_2^{\frac{1}{3}}$
- Start by computing

$$\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_1} = \frac{2}{3}\hat{x}_1^{\frac{-1}{3}}\hat{x}_2^{\frac{1}{3}} = \frac{2}{3}(\frac{\hat{x}_2}{\hat{x}_1})^{\frac{1}{3}},$$

- and

$$\frac{\partial u(\hat{x}_1, \hat{x}_2)}{\partial x_2} = \frac{1}{3}\hat{x}_1^{\frac{2}{3}}\hat{x}_2^{\frac{-2}{3}} = \frac{1}{3}(\frac{\hat{x}_1}{\hat{x}_2})^{\frac{2}{3}}.$$

- Therefore

$$MRS_{x_1,x_2}(\hat{x}_1, \hat{x}_2) = \frac{\frac{2}{3}(\frac{\hat{x}_2}{\hat{x}_1})^{\frac{1}{3}}}{\frac{1}{3}(\frac{\hat{x}_1}{\hat{x}_2})^{\frac{2}{3}}} = 2\frac{\hat{x}_2}{\hat{x}_1} = 1.$$

Exercise: Compute the marginal rate of substitution at the same point for $u(x_1, x_2) = \frac{2}{3}\ln x_1 + \frac{1}{3}\ln x_2$ and compare the results.

Profit maximizing firm

- A firm produces output from labor $l \geq 0$ and capital $k \geq 0$ according to the production function $y = f(k, l)$.
- The profit of a firm at inputs (k, l) can be computed as:

$$g(k, l) = pf(k, l) - rk - wl,$$

- $p > 0$ is the price of the output, $r > 0$ is the rental cost of capital per unit and $w > 0$ is the wage cost of labor per unit.
- Compute the gradient to $g : \mathbb{R}^2_+ \to \mathbb{R}$:

$$\nabla g(k, l) = \begin{pmatrix} p\frac{\partial f(k,l)}{\partial k} - r \\ p\frac{\partial f(k,l)}{\partial l} - w \end{pmatrix}.$$

Profit maximizing firm

- The partial derivative of the production function with respect to k:n is called the *marginal product* of capital $MP_k(k,l)$
- The partial derivative w.r.t. l is called marginal product of labor $MP_l(k,l)$.
- The gradient can therefore be written as:

$$\nabla g(k,l) = \begin{pmatrix} pMP_k(k,l) - r \\ pMP_l(k,l) - w \end{pmatrix}.$$

- We saw before that for if $(\hat{k},\hat{l})$ maximizes profit at $\hat{k},\hat{l} > 0$, then $\nabla g\left(\hat{k},\hat{l}\right) = 0$.
- But this is the case only if

$$\frac{MP_k(\hat{k},\hat{l})}{MP_l(\hat{k},\hat{l})} = \frac{r}{w}.$$

Profit maximizing firm

- We call the ratio $\frac{MP_k(\hat{k},\hat{l})}{MP_l(\hat{k},\hat{l})}$ the *marginal rate of technical substitution between k and l*, $MRTS_{k,l}(\hat{k},\hat{l})$.
- If you give up a small unit of capital at $(\hat{k},\hat{l})$, you need $MRTS_{k,l}(\hat{k},\hat{l})$ additional units of labor to remain at the same level of output
  Exercise Show that the last statement is true by considering the directional derivative in direction $(1, MRTS_{k,l}(\hat{k},\hat{l}))$.
- We call the set of input vectors $\{(k,l)|f(k,l) = \hat{y}\}$ the isoquant of the firm at production level y
- It is the analogue of indifference curves
  Exercise: Compute $MRTS_{k,l}(10, 15)$ for $y = k^{(0.2)}l^{(0.8)}$.

Comparative statics: motivating examples

- Solutions to economic models are often in the form of a system of equations
- We have seen linear models of market equilibrium
- In the previous section, we saw that the solution to profit maximization problem usually takes place at a point where

$$\frac{\partial f(k,l)}{\partial k} - \frac{r}{p} = 0,$$
$$\frac{\partial f(k,l)}{\partial l} - \frac{w}{p} = 0.$$

- Unless f takes a very special for, no explicit solution to the problem is available
- Can we still say do the solutions to this system change when some of the prices p, r, w change?
- You will have a chance to work on this in Problem Set 2

Comparative statics: motivating examples

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

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

  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:

$$Ay + Bx = 0,$$

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

Linear implicit function theorem

- In general form, we then have:

$$f\left(y;x\right)=0.$$

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

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

and consider the effect of a small change $(dy;dx)=(dy_1,...,dy_n;dx_1,...,dx_m)$ on the value of f :

$$f\left(\widehat{y}+dy,\widehat{x}+dx\right)-f\left(\widehat{y},\widehat{x}\right)=Ady+Bdx$$

$$=D_y f\left(\widehat{y};\widehat{x}\right)dy+D_x f\left(\widehat{y};\widehat{x}\right)dx,$$

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

Linear implicit function theorem

- For
$$f\left(y;x\right)=0.$$
to hold at $(y,x)=\left(\widehat{y}+dy,\widehat{x}+dx\right)$, the change must be zero:
$$D_y f(\widehat{y};\widehat{x})dy+D_x f(\widehat{y};\widehat{x})dx=0.$$
- In other words
$$dy=-D_y f(\widehat{y};\widehat{x})^{-1}D_x f(\widehat{y};\widehat{x})dx \quad = \quad A^{-1}Bdx.$$
- If a single exogenous variable changes, then Bdx is a column vector and dy can be solved using Cramer's rule.
- This equation has a solution for all dx only if A^{-1} exists, i.e. if $A=D_y f(\widehat{y};\widehat{x})$ has full rank.
- This result can be generalized for the non-linear case in a neighborhood of $\left(\widehat{y};\widehat{x}\right)$ and it is the implicit function theorem.

Linear implicit function theorem: Example

Consider the system:

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

In matrix form:

$$\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 3 \\ -1 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

or

$$\begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix} x.$$

By Cramer's rule:

$$y_1 = \frac{\det \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix} x}{\det \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}} = \frac{-2}{3} x, \; y_2 = \frac{\det \begin{pmatrix} 2 & -3 \\ 1 & 1 \end{pmatrix} x}{\det \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}} = \frac{-5}{3} x.$$

As you can see, in the linear case, we get an explicit solution.

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 1.
- 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 1.
- 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 $\frac{\partial f(0,1)}{\partial y} \neq 0$.

- The theorem below generalizes the message of this example.
- We say that a function is continuously differentiable at point $(\hat{y}, \hat{x})$ if its partial derivatives w.r.t. x and y exists and are continuous at $(\hat{y}, \hat{x})$.
- An ε -neighborhood $B^{\varepsilon}(\hat{x})$ of $\hat{x} \in \mathbb{R}^k$ is the open set of all points at distance less than ε from $\hat{x}$:

$$B^{\varepsilon}(\hat{x}) := \{x \in \mathbb{R}^k : \|x - \hat{x}\| < \varepsilon\}.$$

One-dimensional implicit function theorem

Theorem

Let $f(y,x)$ be a continuously differentiable at $(\widehat{y},\widehat{x})$ in an ε -neighborhood $B^{\varepsilon}(\widehat{y},\widehat{x})$, for some $\varepsilon > 0$ and also that

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

If $\frac{\partial f(\widehat{y},\widehat{x})}{\partial y} \neq 0$, then there exists a $\delta > 0$ and a continuously differentiable function $y(x)$ in some δ -neighborhood of $\widehat{x}$, $B^{\delta}(\widehat{x})$, such that:

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

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

One-dimensional implicit function theorem: comments

- I have given the theorem in its full mathematical generality.
- The important point to remember from all this is that the key to the theorem is that i) f is continuously differentiable at the solution $(\hat{y}, \hat{x})$, ii) $\frac{\partial f(\hat{y},\hat{x})}{\partial y} \neq 0$.
- Let's write this a bit differently:

$$Df(0,1) = \left(\frac{\partial f(0,1)}{\partial y}, \frac{\partial f(0,1)}{\partial x} \right).$$

For small (dy, dx), we have by the definition of the derivative:

$$f(0+dy, 1+dx) - f(0,1) = \left(\frac{\partial f(0,1)}{\partial y}, \frac{\partial f(0,1)}{\partial x} \right) \begin{pmatrix} dy \\ dx \end{pmatrix}.$$

- For the equation to remain valid at $(0+dy, 1+dx)$,the change in f has to be zero:

$$\frac{\partial f(0,1)}{\partial y} dy + \frac{\partial f(0,1)}{\partial x} dx = 0.$$

One-dimensional implicit function theorem: comments

- By solving for dy as a function of dx, we get:

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

- This approach is the easiest to generalize to get the full implicit function theorem. Notice how nicely this plays on the linearity of the derivative for small changes.

The Implicit function theorem

- Consider now a continuously differentiable non-linear function

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

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

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

- Use the derivative of $Df(\hat{x}, \hat{x})$ to approximate f at $(\hat{y} + dy, \hat{x} + dx)$:

$$f\left(\widehat{y} + dy, \widehat{x} + dx\right) - f\left(\widehat{y}, \widehat{x}\right) = Df\left(\widehat{y}, \widehat{x}\right) = D_y f\left(\widehat{y}; \widehat{x}\right) dy + D_x f\left(\widehat{y}, \widehat{x}\right) dx,$$

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

$$D_y f\left(\widehat{y}, \widehat{x}\right) dy + D_x f\left(\widehat{y}, \widehat{x}\right) dx = 0.$$

- Since $D_y f\left(\widehat{y}; \widehat{x}\right)$ ja $D_x f\left(\widehat{y}; \widehat{x}\right)$ 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 $(\hat{y}, \hat{x})$.

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

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

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

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

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

The Implicit function theorem: An example

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

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

and therefore:

$$dy = \frac{1}{4}\begin{pmatrix} -2 & -2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix} dx.$$

- We could single out e.g. the effect of a change in x_1 on the endogenous variables near $(\hat{y}_1, \hat{y}_2, \hat{x}_1 \hat{x}_2) = (1,1,2,2)$:

$$\begin{pmatrix} \frac{\partial f_1\left(\widehat{y},\widehat{x}\right)}{\partial y_1} & \frac{\partial f_1\left(\widehat{y},\widehat{x}\right)}{\partial y_2} \\ \frac{\partial f_2\left(\widehat{y},\widehat{x}\right)}{\partial y_1} & \frac{\partial f_2\left(\widehat{y},\widehat{x}\right)}{\partial y_2} \end{pmatrix}\begin{pmatrix} dy_1 \\ dy_2 \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1\left(\widehat{y},\widehat{x}\right)}{\partial x_1} \\ \frac{\partial f_2\left(\widehat{y},\widehat{x}\right)}{\partial x_1} \end{pmatrix} dx_1 = 0$$

The Implicit function theorem: An example

- Plugging in $(1, 1, 2, 2)$, we get:

$$\begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} dy_1 \\ dy_2 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \end{pmatrix} dx_1 = 0$$

- Solving by Cramer's rule gives:

$$dy_1 = \frac{\det \begin{pmatrix} 2 & 2 \\ -1 & -2 \end{pmatrix} dx_1}{\det \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix}} = \frac{1}{2} dx_1, \quad dy_2 = \frac{\det \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}}{\det \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix} dx_1} = \frac{3}{4} dx_1.$$

The Implicit function theorem: Main theorem

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

Theorem

Let $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$ *be continuously differentiable in a neighborhood* $B^{\varepsilon}\left(\widehat{y}, \widehat{x}\right)$, *of* $(\hat{y}, \hat{x})$ *for some* $\varepsilon > 0$ *and*

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

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

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

$$Dy\left(\widehat{x}\right) = -\left(D_y f\left(\widehat{y}; \widehat{x}\right)\right)^{-1} D_x f\left(\widehat{y}; \widehat{x}\right)$$

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.