IMPLICIT AND INVERSE FUNCTION THEOREMS

WRITTEN BY PAUL SCHRIMPF AND MODIFIED BY HIRO KASAHARA[1]
NOVEMBER 9, 2020
UNIVERSITY OF BRITISH COLUMBIA
ECONOMICS 526

We have extensively studied how to solve systems of linear equations. We know how to check whether solutions exist and whether they are unique. The inverse and implicit function theorems provide similar results for nonlinear equations.

1. INVERSE FUNCTIONS

Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$. If we know $f(x) = y$, when can we solve for x in terms of y? In other words, when is f invertible? Well, suppose we know that $f(a) = b$ for some $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. Then we can expand f around a,

$$f(x) = f(a) + Df_a(x - a) + r_1(a, x - a)$$

where $r_1(a, x - a)$ is small. Since r_1 is small, we can hopefully ignore it then $y = f(x)$ can be rewritten as a linear equation:

$$\begin{aligned} f(a) + Df_a(x - a) &= y \\ Df_a x &= y - f(a) + Df_a a \end{aligned}$$

we know that this equation has a solution if $\text{rank} Df_a = \text{rank} \begin{pmatrix} Df_a & y - f(a) + Df_a a \end{pmatrix}$. It has a solution for any y if $\text{rank} Df_a = m$. Moreoever, this solution is unique if $\text{rank} Df_a = n$. This discussion is not entirely rigorous because we have not been very careful about what r_1 being small means. The following theorem makes it more precise.

Theorem 1.1 (Inverse function). *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable on an open set E. Let $a \in E$, $f(a) = b$, and Df_a be invertible . Then*

1. *there exist open sets U and V such that $a \in U$, $b \in V$, f is one-to-one on U and $f(U) = V$, and*
2. *the inverse of f exists and is continuously differentiable on V with derivative $\left(Df_{f^{-1}(x)}\right)^{-1}$.*

The open sets U and V are the areas where r_1 is small enough. The continuity of f and its derivative are also needed to ensure that r_1 is small enough. The proof of this theorem is a bit long, but the main idea is the same as the discussion preceding the theorem.

Comment 1.1. The proof uses the fact that the space of all continuous linear transformations between two normed vector spaces is itself a vector space. I do not think we have talked about this before. Anyway, it is a useful fact that already came up in the proof that continuous Gâteaux differentiable implies Fréchet differentiable last lecture. Let V

[1]Thanks to Dana Galizia for corrections.

and W be normed vector spaces with norms $\|\cdot\|_V$ and $\|\cdot\|_W$. Let $BL(V,W)$ denote the set of all continuous (or equivalently bounded) linear transformations from V to W. Then $BL(V,W)$ is a normed vector space with norm

$$\|A\|_{BL} \equiv \sup_{v\in V} \frac{\|Av\|_W}{\|v\|_V}.$$

This is sometimes called the operator norm on $BL(V,W)$. Last lecture, the proof that Gâteaux differentiable implies Fréchet differentiable required that the mapping from V to $BL(V,W)$ defined by Df_x as a function of $x \in V$ had to be continuous with respect to the above norm.

We will often use the inequality,

$$\|Av\|_W \leq \|A\|_{BL} \|v\|_V ,$$

which follows from the definition of $\|\cdot\|_{BL}$. We will also use the fact that if V is finite dimensional and $f(x,v) : V \times V \to W$, is continuous in x and v and linear in v for each x, then $f(x,\cdot) : V \to BL(V,W)$ is continuous in x with respect to $\|\cdot\|_{BL}$.

Proof. For any $y \in \mathbb{R}^n$, consider $\varphi^y(x) = x + Df_a^{-1}(y - f(x))$. By the mean value theorem for $x_1, x_2 \in U$, where $a \in U$ and U is open,

$$\varphi^y(x_1) - \varphi^y(x_2) = D\varphi^y_{\bar{x}}(x_1 - x_2)$$

Note that

$$\begin{aligned} D\varphi^y_{\bar{x}} =& I - Df_a^{-1} Df_{\bar{x}} \\ =& Df_a^{-1}(Df_a - Df_{\bar{x}}). \end{aligned}$$

Since Df_x is continuous (as a function of x) if we make U small enough, then $Df_a - Df_{\bar{x}}$ will be near 0. Let $\lambda = \frac{1}{2\|Df_a^{-1}\|_{BL}}$. Choose U small enough that $\|Df_a - Df_x\| < \lambda$ for all $x \in U$. From above, we know that

$$\begin{aligned} \|\varphi^y(x_1) - \varphi^y(x_2)\| =& \left\| Df_a^{-1}(Df_a - Df_{\bar{x}})(x_1 - x_2) \right\| \\ \leq& \left\| D\varphi^y_x \right\|_{BL} \|Df_a - Df_x\|_{BL} \|x_1 - x_2\| \\ \leq& \frac{1}{2} \|x_1 - x_2\| \end{aligned} \tag{1}$$

For any $y \in f(U)$ we can start with an arbitrary $x_1 \in U$, then create a sequence by setting

$$x_{i+1} = \varphi^y(x_i).$$

From (1), this sequence satisfies

$$\|x_{i+1} - x_i\| \leq \frac{1}{2} \|x_i - x_{i-1}\| .$$

Using this it is easy to verify that x_i form a Cauchy sequence, so it converges. The limit satisfy $\varphi^y(x) = x$, i.e. $f(x) = y$. Moreover, this x is unique because if $\varphi^y(x_1) = x_1$ and

$\varphi^y(x_2) = x_2$, then we have $\|x_1 - x_2\| \leq \frac{1}{2}\|x_1 - x_2\|$, which is only possible if $x_1 = x_2$. [1] Thus for each $y \in f(U)$, there is exactly one x such that $f(x) = y$. That is, f is one-to-one on U. This proves the first part of the theorem and that f^{-1} exists.

We now show that f^{-1} is continuously differentiable with the stated derivative. Let $y, y+k \in V = f(U)$. Then $\exists x, x+h \in U$ such that $y = f(x)$ and $y + k = f(x+h)$. With φ^y as defined above, we have

$$\begin{aligned}\varphi^y(x+h) - \varphi^y(x) =& h + Df_a^{-1}(f(x) - f(x+h)) \\ =& h - Df_a^{-1}k\end{aligned}$$

By 1, $\left\|h - Df_a^{-1}k\right\| \leq \frac{1}{2}\|h\|$. It follows that $\left\|Df_a^{-1}k\right\| \geq \frac{1}{2}\|h\|$ and

$$\|h\| \leq 2\left\|Df_a^{-1}\right\|_{BL}\|k\| = \lambda^{-1}\|k\|.$$

Importantly as $k \to 0$, we also have $h \to 0$. Now,

$$\begin{aligned}\frac{\left\|f^{-1}(y+k) - f^{-1}(y) - Df_x^{-1}k\right\|}{\|k\|} =& \frac{\left\|-Df_x^{-1}(f(x+h) - f(x) - Df_x h)\right\|}{\|k\|} \\ \leq & \|Df_x\|^{-1}\lambda\frac{\|f(x+h) - f(x) - Df_x h\|}{\|h\|} \\ \lim_{k\to 0}\frac{\left\|f^{-1}(y+k) - f^{-1}(y) - Df_x^{-1}k\right\|}{\|k\|} \leq & \lim_{k\to 0}\|Df_x\|_{BL}^{-1}\lambda\frac{\|f(x+h) - f(x) - Df_x h\|}{\|h\|} = 0\end{aligned}$$

Finally, since Df_x is continuous, so is $(Df_{f^{-1}(y)})^{-1}$, which is the derivative of f^{-1}. □

The proof of the inverse function theorem might be a bit confusing. The important idea is that if the derivative of a function is nonsingular at a point, then you can invert the function around that point because inverting the system of linear equations given by the mean value expansion around that point nearly gives the inverse of the function.

2. Implicit functions

The implicit function theorem is a generalization of the inverse function theorem. In economics, we usually have some variables, say x, that we want to solve for in terms of some parameters, say β. For example, x could be a person's consumption of a bundle of goods, and b could be the prices of each good and the parameters of the utility function. Sometimes, we might be able to separate x and β so that we can write the conditions of our model as $f(x) = b(\beta)$. Then we can use the inverse function theorem to compute $\frac{\partial x_i}{\partial \beta_j}$ and other quantities of interest. However, it is not always easy and sometimes not possible to separate x and β onto opposite sides of the equation. In this case our model gives us equations of the form $f(x, \beta) = c$. The implicit function theorem tells us when we can solve for x in terms of β and what $\frac{\partial x_i}{\partial \beta_j}$ will be.

[1] Functions like φ^y that have $d(\phi(x), \phi(y)) \leq c d(x,y)$ for $c < 1$ are called contraction mappings. The x with $x = \phi(x)$ is called a fixed point of the contraction mapping. The argument in the proof shows that contraction mappings have at most one fixed point. It is not hard to show that contraction mappings always have exactly one fixed point.

The basic idea of the implicit function theorem is the same as that for the inverse function theorem. We will take a first order expansion of f and look at a linear system whose coefficients are the first derivatives of f. Let $f : \mathbb{R}^n \to \mathbb{R}^m$. Suppose f can be written as $f(x,y)$ with $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^{n-k}$. x are endogenous variables that we want to solve for, and y are exogenous parameters. We have a model that requires $f(x,y) = c$, and we know that some particular x_0 and y_0 satisfy $f(x_0,y_0) = c$. To solve for x in terms of y, we can expand f around x_0 and y_0.

$$f(x,y) = f(x_0,y_0) + D_x f_{(x_0,y_0)}(x - x_0) + D_y f_{(x_0,y_0)}(y - y_0) + r(x,y) = c$$

In this equation, $D_x f_{(x_0,y_0)}$ is the m by k matrix of first partial derivatives of f with respect to x evaluated at (x_0,y_0). Similary, $D_y f_{(x_0,y_0)}$ is the m by $n-k$ matrix of first partial derivatives of f with respect to y evaluated at (x_0,y_0). Then, if $r(x,y)$ is small enough, we have

$$\begin{aligned} f(x_0,y_0) + D_x f_{(x_0,y_0)}(x - x_0) + D_y f_{(x_0,y_0)}(y - y_0) &\approx c \\ D_x f_{(x_0,y_0)}(x - x_0) &\approx \left(c - f(x_0,y_0) - D_y f_{(x_0,y_0)}(y - y_0)\right) \end{aligned}$$

This is just a system of linear equations with unknowns $(x - x_0)$. If $k = m$ and $D_x f_{(x_0,y_0)}$ is nonsingular, then we have

$$x \approx x_0 + \left(D_x f_{(x_0,y_0)}\right)^{-1} \left(c - f(x_0,y_0) - D_y f_{(x_0,y_0)}(y - y_0)\right)$$

which gives x approximately as function of y. The implicit function says that you can make this approximation exact and get $x = g(y)$. The theorem also tells you what the derivative of $g(y)$ is in terms of the derivative of f.

Theorem 2.1 (Implicit function)**.** *Let $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$ be continuously differentiable on some open set E and suppose $f(x_0,y_0) = c$ for some $(x_0,y_0) \in E$, where $x_0 \in \mathbb{R}^n$ and $y_0 \in \mathbb{R}^m$. If $D_x f_{(x_0,y_0)}$ is invertible, then there exists open sets $U \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^{n-k}$ with $x_0 \in U$ and $y_0 \in W$ such that*

(1) *For each $y \in W$ there is a unique x such that $(x,y) \in U$ and $f(x,y) = c$.*
(2) *Define this x as $g(y)$. Then g is continuously differentiable on W, $g(y_0) = x_0$, $f(g(y),y) = c$ for all $y \in W$, and $Dg_{y_0} = -\left(D_x f_{(x_0,y_0)}\right)^{-1} D_y f_{(x_0,y_0)}$*

Proof. We will show the first part by applying the inverse function theorem. Define $F : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$ by $F(x,y) = (f(x,y),y)$. To apply the inverse function theorem we must show that F is continuously differentiable and $DF_{(x_0,y_0)}$ is invertible. To show that F is continuously differentiable, note that

$$\begin{aligned} F(x+h,y+k) - F(x,y) &= (f(x+h,y+k) - f(x,y),k) \\ &= (Df_{(\bar{x},\bar{y})}(h\,k),k) \end{aligned}$$

where the second line follows from the mean value theorem. It is then apparent that

$$\lim_{(h,k)\to 0} \frac{\left\| F(x+h,y+k) - F(x,y) - \begin{pmatrix} D_x f_{(x,y)} & D_y f_{(x,y)} \\ 0 & I_m \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \right\|}{\|(h,k)\|} = 0.$$

So, $DF_{(x,y)} = \begin{pmatrix} D_x f_{(x,y)} & D_y f_{(x,y)} \\ 0 & I_m \end{pmatrix}$, which is continuous sinve $Df_{(x,y)}$ is continuous. Also, $DF_{(x_0,y_0)}$ can be shown to be invertible by using the partitioned inverse formula because $D_x f_{(x_0,y_0)}$ is invertiable by assumption. Therefore, by the inverse function theorem, there exists open sets U and V such that $(x_0, y_0) \in U$ and $(c, y_0) \in V$, and F is one-to-one on U.

Let W be the set of $y \in \mathbb{R}^m$ such that $(c, y) \in V$. By definition, $y_0 \in W$. Also, W is open in $\mathbb{R}^m$ because V is open in $\mathbb{R}^{n+m}$.

We can now complete the proof of 1. If $y \in W$ then $(c, y) = F(x, y)$ for some $(x, y) \in U$. If there is another (x', y) such that $f(x', y) = c$, then $F(x', y) = (c, y) = F(x, y)$. We just showed that F is one-to-one on U, so $x' = x$.

We now prove 2. Define $g(y)$ for $y \in W$ such that $(g(y), y) \in U$ and $f(g(y), y) = c$, and

$$F(g(y), y) = (c, y).$$

By the inverse function theorem, F has an inverse on U. Call it G. Then

$$G(c, y) = (g(y), y)$$

and G is continuously differentiable, so g must be as well. Differentiating the above equation with respect to y, we have

$$D_y G_{(c,y)} = \begin{pmatrix} Dg_y \\ I_m \end{pmatrix}$$

On the other hand, from the inverse function theorem, the derivative of G at (x_0, y_0) is

$$\begin{aligned} DG_{(x_0,y_0)} &= \left(DF_{(x_0,y_0)} \right)^{-1} \\ &= \begin{pmatrix} D_x f_{(x_0,y_0)} & D_y f_{(x_0,y_0)} \\ 0 & I_m \end{pmatrix}^{-1} \\ &= \begin{pmatrix} D_x f^{-1}_{(x_0,y_0)} & -D_x f^{-1}_{(x_0,y_0)} D_y f_{(x_0,y_0)} \\ 0 & I_m \end{pmatrix} \end{aligned}$$

In particular,

$$D_y G_{(c,y_0)} = \begin{pmatrix} -D_x f^{-1}_{(x_0,y_0)} D_y f_{(x_0,y_0)} \\ I_m \end{pmatrix} = \begin{pmatrix} Dg_{y_0} \\ I_m \end{pmatrix}$$

so $Dg_{y_0} = -D_x f^{-1}_{(x_0,y_0)} D_y f_{(x_0,y_0)}$. □

3. Contraction mappings

One step of the proof the of the inverse function theorem was to show that

$$\|\varphi^y(x_1) - \varphi^y(x_2)\| \leq \frac{1}{2} \|x_1 - x_2\|.$$

This property ensures that $\varphi(x) = x$ has a unique solution. Functions like φ^y appear quite often, so they have name.

Definition 3.1. Let $f : \mathbb{R}^n \to \mathbb{R}^n$. f is a **contraction mapping** on $U \subseteq \mathbb{R}^n$ if for all $x, y \in U$,

$$\|f(x) - f(y)\| \leq c \|x - y\|$$

for some $0 \leq c < 1$.

If f is a contraction mapping, then an x such that $f(x) = x$ is called a **fixed point** of the contraction mapping. Any contraction mapping has at most one fixed point.

Lemma 3.1. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a contraction mapping on $U \subseteq \mathbb{R}^n$. If $x_1 = f(x_1)$ and $x_2 = f(x_2)$ for some $x_1, x_2 \in U$, then $x_1 = x_2$.*

Proof. Since f is a contraction mapping,

$$\|f(x_1) - f(x_2)\| \leq c \|x_1 - x_2\| .$$

$f(x_i) = x_i$, so

$$\|x_1 - x_2\| \leq c \|x_1 - x_2\| .$$

Since $0 \geq c < 1$, the previous inequality can only be true if $\|x_1 - x_2\| = 0$. Thus, $x_1 = x_2$. □

Starting from any x_0, we can construct a sequence, $x_1 = f(x_0)$, $x_2 = f(x_1)$, etc. When f is a contraction, $\|x_n - x_{n+1}\| \leq c^n \|x_1 - x_0\|$, which approaches 0 as $n \to \infty$. Thus, $\{x_n\}$ is a Cauchy sequence and converges to a limit. Moreover, this limit will be such that $x = f(x)$, i.e. it will be a fixed point.

Lemma 3.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a contraction mapping on $U \subseteq \mathbb{R}^n$, and suppose that $f(U) \subseteq U$. Then f has a unique fixed point.*

Proof. Pick $x_0 \in U$. As in the discussion before the lemma, construct the sequence defined by $x_n = f(x_{n-1})$. Each $x_n \in U$ because $x_n = f(x_{n-1}) \in f(U)$ and $f(U) \subseteq U$ by assumption. Since f is a contraction on U, $\|x_{n+1} - x_n\| \leq c^n \|x_1 - x_0\|$, so $\lim_{n\to\infty} \|x_{n+1} - x_n\| = 0$, and $\{x_n\}$ is a Cauchy sequence. Let $x = \lim_{n\to\infty} x_n$. Then

$$\begin{aligned} \|x - f(x)\| &\leq \|x - x_n\| + \|f(x) - f(x_{n-1})\| \\ &\leq \|x - x_n\| + c \|x - x_{n-1}\| \end{aligned}$$

$x_n \to x$, so for any $\epsilon > 0$ $\exists N$, such that if $n \geq N$, then $\|x - x_n\| < \frac{\epsilon}{1+c}$. Then,

$$\|x - f(x)\| < \epsilon$$

for any $\epsilon > 0$. Therefore, $x = f(x)$. □

4. Applications

This lecture and the previous one have been rather theoretical, so this section goes over a couple of applications of what has been covered.

4.1. **Roy's Identity.** Let $V(m, p)$ be an **indirect utility function**. Given total expenditure m and a vector of prices p, the maximum utility that a person can achieve is $V(m, p)$. If U is the utility function, the indirect utility function is given by

$$V(m, p) = \max_c U(c) \text{ s.t. } pc \leq m. \tag{2}$$

Similarly, **expenditure function**, $E(u, p)$, is the minimum amount of money that can be spent to achieve utility u when faced with prices p. That is,

$$E(u, p) = \min_c pc \text{ s.t. } U(c) \geq u. \tag{3}$$

We haven't yet covered optimization, so let's just assume that (2) and (3) have unique solutions. In normal cases, we would expect that $V(E(u, p), p) = u$ and $E(V(m, p), p) = m$. Let's come up with conditions that ensure these two equalities hold. Let's start by working with $V(E(u, p), p) = u$. By definition of $E(u, p)$, there must be some c^* such that $pc^* = E(u, p)$ and $U(c^*) = u$. Using that same c^* in (2), we see that $V(E(u, p), p) \geq U(c^*) = u$. Suppose it were strictly greater. Then there is some $\tilde{c}$ such that $U(\tilde{c}) > u$ and $p\tilde{c} \leq pc^* = m$. But if U is continuous, then for any ϵ, we can find $\delta > 0$ such that if $\|h\| < \delta$ then $|U(\tilde{c}) - U(\tilde{c} + h)| < \epsilon$ and in particular, $U(\tilde{c} + h) > u$. If $p \neq 0$, we can choose an h with $\|h\| < \delta$ and $ph < 0$. However, then $p(\tilde{c} + h) < pc^*$, which should not be possible given how we have defined c^*. Thus, assuming U is continuous and $p \neq 0$ is enough to ensure that $V(E(u, p), p) = u$.

Having established, $V(E(u, p), p) = u$, we can differentiate with respect to the price of i good to get

$$\begin{aligned} \frac{\partial V}{\partial m}(E(u, p), p)\frac{\partial E}{\partial p_i}(u, p) + \frac{\partial V}{\partial p_i}(E(u, p), p) =& 0 \\ \frac{\partial E}{\partial p_i}(u, p) =& -\frac{\frac{\partial V}{\partial p_i}(E(u, p), p)}{\frac{\partial V}{\partial m}(E(u, p), p)} \end{aligned}$$

This result combined with **Shephard's lemma** (which we will prove after studying optimization) gives **Roy's identity**. Shephard's lemma is

$$c_i^*(u, p) = \frac{\partial E}{\partial p_i}(u, p),$$

and Roy's identity is

$$c_i^*(m, p) = -\frac{\frac{\partial V}{\partial p_i}(m, p)}{\frac{\partial V}{\partial m}(m, p)}.$$

Roy's identity is very useful because it relates demand, something that we can observe, to the indirect utility function, something that cannot be directly observed.

4.2. **Comparative statics.** Consider a simple finite horizon macro model. The production function is Cobb-Douglas with only one input, capital k_t, so output at time t is

$$y_t = A_t k_t^{\alpha}$$

where A_t is productivity. Output can be either consumed or saved as capital for the next period. Capital depreciates at rate δ. The budget constraint each period is

$$c_t + k_{t+1} = (1-\delta)k_t + A_t k_t^\alpha.$$

The consumer has a standard CRRA utility function that is additively separable over time and discount rate β. The social planner's problem is to maximize total utility subject to the budget constraints,

$$\max_{\{c_t,k_t\}_{t=0}^T} \sum_{t=0}^T \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} \text{ s.t. } c_t + k_{t+1} = (1-\delta)k_t + A_t k_t^\alpha.$$

You are likely already familiar with using Lagrangians to solve constrained maximization problems. If not, do not worry because we will cover it in some detail in a week or two. Anyway, the idea is that we can replace the constrained problem with an unconstrained one of the form:

$$\max_{\{c_t,k_t,\lambda_t\}_{t=0}^T} \sum_{t=0}^T \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} + \lambda_t(c_t + k_{t+1} - (1-\delta)k_t - A_t k_t^\alpha)$$

where we have adding λ_t times the tth constraint to the objective function and we are now maximizing over λ_t as well as c_t and k_t. We have already shown that for c_t, k_t, λ_t, to be local maxima, we must have the derivative of the objective function equal to zero. This is called the first order condition. Here, the first order conditions are

$$\begin{aligned}
[c_t]: \quad && \beta^t c_t^{-\gamma} =& \lambda_t \\
[k_t]: \quad && \lambda_{t-1} =& \lambda_t \left((1-\delta) + A_t \alpha k_t^{\alpha-1} \right) \\
[\lambda_t]: \quad && c_t + k_{t+1} =& (1-\delta)k_t + A_t k_t^\alpha
\end{aligned}$$

The values of c_t, k_t, λ_t that solve the maximization problem necessarily solve the first order conditions as well, but not every solution to the first order conditions solves the maximization problem. There is a second condition that ensures a solution to the first order condition solves the maximization problem. We will overlook second order conditions for now, but we will study them soon. We can eliminate λ_t from the system of first order conditions by combining $[c_t]$ and $[c_t]$ to get

$$\begin{aligned}
\beta^{t-1} c_{t-1}^{-\gamma} =& \beta^t c_t^{-\gamma} \left((1-\delta) + A_t \alpha k_t^{\alpha-1} \right) \\
\left(\frac{c_t}{c_{t-1}} \right)^\gamma =& \beta \left((1-\delta) + A_t \alpha k_t^{\alpha-1} \right)
\end{aligned}$$

Combined with the budget constraint we now have two nonlinear equation for each t with two unknowns for each t. The solution to these equations is the optimal sequence of c_t and k_t.

$$\begin{aligned}
\left(\frac{c_t}{c_{t-1}} \right)^\gamma =& \beta \left((1-\delta) + A_t \alpha k_t^{\alpha-1} \right) \\
c_t + k_{t+1} =& (1-\delta)k_t + A_t k_t^\alpha
\end{aligned}$$

Unfortunately, there is no closed form solution to these equations. However, we can still calculate certain quantities of interest by using the implicit function theorem. For example, what is the effect of changes in productivity, A_t, on consumption, capital, and welfare? Suppose A_t changes unexpectedly at time $T-1$ for one period only, so we can treat c_{T-2} and k_{T-1} as constant. We want to find the change in c_{T-1}, c_T, and k_T. Note that the equations above hold for each t. The relevant three equations here are

$$\begin{aligned} 0 =& F(c_T, c_{T-1}, k_T, A_T, A_{T-1}, c_{T-2}, k_{T-1}) \\ =& \begin{pmatrix} c_{T-1} + k_T - (1-\delta)k_{T-1} - A_{T-1}k_{T-1}^{\alpha} \\ c_T - (1-\delta)k_T - A_T k_T^{\alpha} \\ c_{T-1}^{-\gamma} - c_T^{-\gamma}\beta\left((1-\delta) + A_T\alpha k_T^{\alpha-1}\right) \end{pmatrix} \end{aligned}$$

The implicit function theorem says that

$$\begin{aligned} \begin{pmatrix} \frac{\partial c_{T-1}}{\partial A_{T-1}} \\ \frac{\partial c_T}{\partial A_{T-1}} \\ \frac{\partial k_T}{\partial A_{T-1}} \end{pmatrix} =& - \begin{pmatrix} \frac{\partial F_1}{\partial c_{T-1}} & \frac{\partial F_1}{\partial c_T} & \frac{\partial F_1}{\partial k_T} \\ \frac{\partial F_2}{\partial c_{T-1}} & \frac{\partial F_2}{\partial c_T} & \frac{\partial F_2}{\partial k_T} \\ \frac{\partial F_3}{\partial c_{T-1}} & \frac{\partial F_3}{\partial c_T} & \frac{\partial F_3}{\partial k_T} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial A_{T-1}} \\ \frac{\partial F_2}{\partial A_{T-1}} \\ \frac{\partial F_3}{\partial A_{T-1}} \end{pmatrix} \\ =& - \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -(1-\delta) - A_T\alpha k_T^{\alpha-1} \\ -\gamma c_{T-1}^{-\gamma-1} & \gamma c_T^{-\gamma-1}\beta\left((1-\delta) + A_T\alpha k_T^{\alpha-1}\right) & -c_T^{-\gamma}\beta A_T\alpha(\alpha-1)k_T^{\alpha-2} \end{pmatrix}^{-1} \begin{pmatrix} -k_{T-1}^{\alpha} \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

We can invert this matrix using Gaussian elimination:

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & 1 & k_{T-1}^{\alpha} \\ 0 & 1 & -(1-\delta) - A_T\alpha k_T^{\alpha-1} & 0 \\ -\gamma c_{T-1}^{-\gamma-1} & \gamma c_T^{-\gamma-1}\beta\left((1-\delta) + A_T\alpha k_T^{\alpha-1}\right) & -c_T^{-\gamma}\beta A_T\alpha(\alpha-1)k_T^{\alpha-2} & 0 \end{pmatrix} \simeq \\ \simeq & \begin{pmatrix} 1 & 0 & 1 & k_{T-1}^{\alpha} \\ 0 & 1 & -(1-\delta) - A_T\alpha k_T^{\alpha-1} & 0 \\ 0 & \gamma c_T^{-\gamma-1}\beta\left((1-\delta) + A_T\alpha k_T^{\alpha-1}\right) & -c_T^{-\gamma}\beta A_T\alpha(\alpha-1)k_T^{\alpha-2} + \gamma c_{T-1}^{-\gamma-1} & \gamma c_{t-1}^{-\gamma-1}k_{T-1}^{\alpha} \end{pmatrix} \\ \simeq & \begin{pmatrix} 1 & 0 & 1 & k_{T-1}^{\alpha} \\ 0 & 1 & -(1-\delta) - A_T\alpha k_T^{\alpha-1} & 0 \\ 0 & 0 & E & \gamma c_{t-1}^{-\gamma-1}k_{T-1}^{\alpha} \end{pmatrix} \end{aligned}$$

where

$$\begin{aligned} E =& \left(\gamma c_T^{-\gamma-1}\beta\left((1-\delta) + A_T\alpha k_T^{\alpha-1}\right)\right)\left((1-\delta) + A_T\alpha k_T^{\alpha-1}\right) + \\ & - c_T^{-\gamma}\beta A_T\alpha(\alpha-1)k_T^{\alpha-2} + \gamma c_{T-1}^{-\gamma-1}. \end{aligned}$$

This expression is quite complicated, but we can still make a few observations. First, we generally assume that $\alpha \leq 1$, which ensures that $E > 0$. Then,

$$\frac{\partial k_T}{\partial A_{T-1}} = \frac{\gamma c_{T-1}^{-\gamma-1} k_{T-1}^{\alpha}}{E} > 0$$

So when productivity goes up at time $T-1$, capital at time T increases. Also, from the second equation,

$$\frac{\partial c_T}{\partial A_{T-1}} = \frac{\partial k_T}{\partial A_{T-1}} \left((1-\delta) + A_T \alpha_T k_T^{\alpha-1} \right),$$

so consumption at time T also increases. From the first equation,

$$\begin{aligned}
\frac{\partial c_{T-1}}{\partial A_{T-1}} =& k_{T-1}^{\alpha} - \frac{\partial k_T}{\partial A_{T-1}} \\
=& \frac{k_{T-1}^{\alpha} E - \gamma c_{T-1}^{-\gamma-1} k_{T-1}^{\alpha}}{E} \\
=& \frac{k_{T-1}^{\alpha} \left(\gamma c_T^{-\gamma-1} \beta \left((1-\delta) + A_T \alpha k_T^{\alpha-1} \right) \right) \left((1-\delta) + A_T \alpha k_T^{\alpha-1} \right) - c_T^{-\gamma} \beta A_T \alpha (\alpha-1) k_T^{\alpha-2}}{E} \\
0 \leq & \frac{\partial c_{T-1}}{\partial A_{T-1}} < k_{T-1}^{\alpha}
\end{aligned}$$

So c_{T-1} increases when A_{T-1} increases, but less than the increase in output at time $T-1$.