MA2AA1 (ODE’s): The inverse and implicit function theorem

Sebastian van Strien (Imperial College)

February 3, 2013
Some of you did not do multivariable calculus. This note provides a crash course on this topic. This note also includes some important theorems which are not covered in 2nd year courses.

These notes include examples that are taken from the internet.
Definition of Jacobian

Suppose that $F: U \to V$ where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^p$. We say that $F$ is differentiable at $x \in U$ if there exists a linear map $A: \mathbb{R}^n \to \mathbb{R}^m$ (i.e. a $m \times n$ matrix $A$)

$$ \frac{|(F(x + u) - F(x)) - Au|}{|u|} \to 0 $$

as $u \to 0$. In this case we define $DF_x = A$.

- In other words $F(x + u) = F(x) + Au + o(|u|)$. ($A$ is the linear part of the Taylor expansion of $F$).
- How to compute $DF_x$? This is just the Jacobian matrix, see next slide.
- If $f: \mathbb{R}^n \to \mathbb{R}$ then $Df_x$ is a $1 \times n$ matrix which is also called $\text{grad}(f)$ or $\nabla f(x)$.
Example 1. Let $F(x, y) = \begin{pmatrix} x^2 + yx \\ xy - y \end{pmatrix}$ then

$DF_{x,y} = \begin{pmatrix} 2x + y & x \\ y & x - 1 \end{pmatrix}$.

$(Df_0)u$ is the directional derivative of $f$ (in the direction $u$).

Example 2. If $F(x, y) = \begin{pmatrix} x^2 + yx \\ xy - y \end{pmatrix}$ and $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ then

$(DF_{x,y})e_1 = \begin{pmatrix} 2x + y & x \\ y & x - 1 \end{pmatrix} e_1 = \begin{pmatrix} 2x + y \\ y \end{pmatrix}$. This is what you get when you fix $y$ and differentiate w.r.t. $x$ in $F(x, y)$.

For each fixed $y$ one has a curve $x \mapsto F(x, y) = \begin{pmatrix} x^2 + yx \\ xy - y \end{pmatrix}$ and $(DF_{x,y})e_1 = \begin{pmatrix} 2x + y \\ y \end{pmatrix}$ gives its speed vector.

Remark: Sometimes one writes $DF(x, y)u$ instead of $DF_{x,y}u$. If $u$ is the $i$-th unit vector $e_i$ then one often writes $D_i F_{x,y}$ and if $i = 1$ something like $D_x F(x, y)$. 

Differential Equations MA2AA1
Sebastian van Strien (Imperial College)
Theorem 1. **Multivariable Mean Value Theorem**

If \( f : \mathbb{R} \to \mathbb{R}^m \) is differentiable then for each \( x, y \in \mathbb{R} \) there exists \( \xi \in [x, y] \) so that \( |f(x) - f(y)| \leq |Df_\xi||x - y| \).

**Proof:** By the Main Theorem of integration,
\[
f(y) - f(x) = \int_x^y Df_s \, ds \text{ (where } Df_t \text{ is the } n \times 1 \text{ matrix (i.e. vertical vector) of derivatives of each component of } f). \]

So
\[
|f(x) - f(y)| = |\int_x^y Df_s \, ds| \leq \int_x^y |Df_s| \, ds \leq \max_{s \in (x,y)} |Df_s| |x - y| \leq |Df_\xi||x - y|
\]

for some \( \xi \in [x, y] \).

**Corollary:** If \( f : \mathbb{R}^n \to \mathbb{R}^m \) is continuously differentiable then for each \( x, y \in \mathbb{R}^n \) there exists \( \xi \) in the arc \([x, y]\) connecting \( x \) and \( y \) so that \( |f(x) - f(y)| \leq |Df_\xi(u)||x - y| \) where
\[
u = (x - y)/|x - y|.
\]

**Proof:** just consider \( f \) restricted to the line connecting \( x, y \) and apply the previous theorem.
The inverse function theorem

Theorem 2. Let $U \subset \mathbb{R}^n$ be open, $p \in U$ and $F: U \rightarrow \mathbb{R}^n$ be continuously differentiable and suppose that the matrix $DF_p$ is invertible. Then there exist open sets $W \subset U$ and $V \subset \mathbb{R}^n$ with $p \in W$ and $F(p) \in V$, so that $F: W \rightarrow V$ is a bijection and so that its inverse $G: V \rightarrow W$ is also differentiable.

Definition A map $F: U \rightarrow V$ which has a differentiable inverse is called a \textit{diffeomorphism}.

Proof: Without loss of generality we can assume that $p = 0 = F(p)$ (just apply a translation). By composing with a linear transformation we can even also assume $DF_0 = I$. Since we assume that $x \mapsto DF_x$ is continuous, there exists $\delta > 0$ so that

$$|| I - DF_x || \leq 1/2 \text{ for all } x \in \mathbb{R} \text{ with } |x| \leq 2\delta.$$ (1)

Here, as usual, we define the norm of a matrix $A$ to be

$$|| A || = \sup\{|Ax|; |x| = 1\}.$$
Given $y$ with $|y| \leq \delta/2$ define the transformation

$$T_y(x) = y + x - F(x).$$

Note that

$$T_y(x) = x \iff F(x) = y.$$

So finding a fixed point of $T_y$ gives us a point $x$ for which $G(y) = x$.

We will find $x$ using the Banach Contraction Mapping Theorem.

(Step 1) By (1) we had $\|I - DF_x\| \leq 1/2$ when $|x| \leq 2\delta$.

Therefore, the Mean Value Theorem applied to $x \mapsto x - F(x)$ gives

$$|x - F(x) - (0 - F(0))| \leq \frac{1}{2}|x - 0| \text{ for } |x| \leq 2\delta$$

Therefore if $|x| \leq \delta$ then

$$|T_y(x)| \leq |y| + |x - F(x)| \leq \delta/2 + \delta/2 = \delta.$$
(Step 2) \( T_y : B \to B \) is a contraction since if \( x, z \in B_\delta(0) \) then \( |x - z| \leq 2\delta \) and so we obtain by the Mean Value Theorem again

\[
|T_y(x) - T_y(z)| = |x - F(x) - (z - F(z))| \leq \frac{1}{2} |x - z|. \tag{2}
\]

(Step 3) Since \( B_\delta(0) \) is a Banach space, there exists a unique \( x \in B_\delta(0) \) with \( T_y(x) = x \) i.e. so that \( F(x) = y \).

(Step 4) The upshot is that for each \( y \in B_{\delta/2}(0) \) there is precisely one solution \( x \in B_\delta(0) \) of the equation \( F(x) = y \). Hence there exists \( W \subset B_\delta(0) \) so that the map

\[
F : W \to V := B_{\delta/2}(0)
\]

is a bijection. So \( F : W \to V \) has an inverse, denoted by \( G \).
(Step 5) \( G \) is continuous: Set \( u = F(x) \) and \( v = F(z) \).

Applying the triangle and the 2nd inequality in equation (2),

\[
|x - z| = |(x - z) - (F(x) - F(z)) + (F(x) - F(z))| \leq (x - z) - (F(x) - F(z))| + |F(x) - F(z)| \leq \frac{1}{2}|x - z| + |F(x) - F(z)|.
\]

So \( |G(u) - G(v)| = |x - z| \leq 2|F(x) - F(z)| = 2|u - v| \).

(Step 6) \( G \) is differentiable:

\[
|(G(u) - G(v)) - (DF_z)^{-1}(u - v)| = |x - z - (DF_z)^{-1}(F(x) - F(z))| \leq \|(DF_z)^{-1}\| \cdot |DF_z(x - z) - (F(x) - F(z))| = o(|x - z|) = 2o(|u - v|).
\]

as \( \|(DF_z)^{-1}\| \) is bounded, using the definition and the last inequality in step 5. Hence

\[
|G(u) - G(v) - (DF_z)^{-1}(u - v)| = o(|u - v|)
\]

proving that \( G \) is differentiable and that \( DG_v = (DF_z)^{-1} \).
Example 3. Consider the set of equations
\[
x^2 + \frac{y^2}{x} = u, \sin(x) + \cos(y) = v.
\]
Given \((u, v)\) near \((u_0, v_0) = (2, \cos(1) + \sin(1))\) is it possible to find a unique \((x, y)\) near to \((x_0, y_0) = (1, 1)\) satisfying this set of equations? To check this, we define
\[
F(x, y) = \begin{pmatrix}
\frac{x^2+y^2}{x} \\
\sin(x) + \cos(y)
\end{pmatrix}.
\]

The Jacobian matrix is
\[
\begin{pmatrix}
\frac{x^2-y^2}{x^2} & \frac{2y}{x} \\
\cos(x) & -\sin(y)
\end{pmatrix}.
\]

The determinant of this is \(\frac{y^2-x^2}{x^2} \sin(y) - \frac{2y}{x} \cos(x)\) which is non-zero near \((1, 1)\). So for every \((u, v)\) sufficiently close to \((u_0, v_0)\) one can find a unique solution near to \((x_0, y_0)\) to this set of equations. Near \((\pi/2, \pi/2)\) probably not.
Theorem 3. **Implicit Function Theorem** Let \( F: \mathbb{R}^p \times \mathbb{R}^n \rightarrow \mathbb{R}^n \) be differentiable and assume that \( F(0, 0) = 0 \).

Moreover, assume that \( n \times n \) matrix obtained by deleting the first \( p \) columns of the matrix \( DF_{0,0} \) is invertible.

Then there exists a function \( G: \mathbb{R}^p \rightarrow \mathbb{R}^n \) so that for all \((x, y)\) near \((0, 0)\)

\[
y = G(x) \iff F(x, y) = 0.
\]

The proof is a fairly simple application of the inverse function theorem, and won’t be given here. The \( \mathbb{R}^p \) part in \( \mathbb{R}^p \times \mathbb{R}^n \) can be thought as parameters.
Example 4. Let \( f(x, y) = x^2 + y^2 - 1 \). Then one can consider this as locally as a function \( y(x) \) when \( \partial f/\partial y = 2y \neq 0 \).

Example 5. What can you say about solving the equations

\[
\begin{align*}
x^2 - y^2 - u^3 + v^2 + 4 &= 0 \\
2xy + y^2 - 2u^2 + 3v^4 + 8 &= 0
\end{align*}
\]

for \( u, v \) in terms of \( x, y \) in a neighbourhood of the solution \( (x, y, y, v) = (2, -1, 2, 1) \). Define

\[ F(x, y, y, v) = (x^2 - y^2 - u^3 + v^2 + 4, 2xy + y^2 - 2u^2 + 3v^4 + 8). \]

We have to consider the part of the Jacobian matrix which concerns the derivatives w.r.t. \( u, v \) at this point. That is

\[
\begin{pmatrix}
-3u^2 & 2v \\
-4u & 12v^3
\end{pmatrix}
\bigg|_{(2, -1, 2, 1)} =
\begin{pmatrix}
-12 & 2 \\
-8 & 12
\end{pmatrix}
\]

which is an invertible matrix.

So locally, near \( (2, -1, 2, 1) \) one can write

\[(u, v) = G(x, y) \text{ that is } F(x, y, G_1(x, y), G_2(x, y)) = 0.\]
To determine \( \partial G_1/\partial x \) (i.e. \( \partial u/\partial x \)) we differentiate this. Indeed, writing \( u = G(x, y) \) and \( v = G(x, y) \) and differentiating
\[
\begin{align*}
  x^2 - y^2 - u^3 + v^2 + 4 &= 0, \\
  2xy + y^2 - 2u^2 + 3v^4 + 8 &= 0,
\end{align*}
\]
w.r.t. \( x \) one has
\[
\begin{align*}
  2x - 3u^2 \frac{\partial u}{\partial x} + 2v \frac{\partial v}{\partial x} &= 0, \\
  2y - 4u \frac{\partial u}{\partial x} + 12v^3 \frac{\partial v}{\partial x} &= 0.
\end{align*}
\]
So
\[
\left( \begin{array}{c} \frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial x} \end{array} \right) = \left( \begin{array}{cc} 3u^2 & -2v \\ 4u & -12v^3 \end{array} \right) \left( \begin{array}{c} 2x \\ 2y \end{array} \right) = \frac{1}{8uv - 36u^2v^2} \left( \begin{array}{cc} -12v^3 & 2v \\ -4u & 3u^2 \end{array} \right) \left( \begin{array}{c} 2x \\ 2y \end{array} \right)
\]
Hence
\[
\frac{\partial u}{\partial x} = \frac{-24xv^3 + 4vy}{8uv - 36u^2v^2}.
\]