Multivariable Calculus Winter Notes

Camila Restrepo

Last updated January 23, 2024

Note: Theorem numbers come from the order they are presented in lecture, and do not correspond to any textbook or written course material.

Week 1

Classifying Critical Points

Theorem 1.1 2nd Derivative Test
Let $f \in C^2(\Omega)$ and let $a \in \Omega(\Omega \subseteq \mathbb{R}^n)$ be a critical point of f.

1. If $H_f(a)$ is positive definite then f has a local minimum at a.
2. If $H_f(a)$ is negative definite then f has a local maxiumum at a.
3. If $H_f(a)$ is indefinite then f has a saddle point at a.

Recall: Any symmetric $n \times n$ matrix A can be diagonalized, i.e., $\exists$ an orthonormal basis $u_1, u_2, \ldots, u_n$ in $\mathbb{R}^n$ and real numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $Au_i = \lambda_i u_i \forall i = 1, 2, \ldots, n$.

Proposition 1.2
Let Q be the quadratic form associated with an $n \times n$ symmetric matrix A. Then:

1. Q is positive $\iff$ all the eigenvalues of A are positive,
2. Q is negative $\iff$ all the eigenvalues of A are negative,
3. Q is indefinite $\iff$ A has both positive and negative eigenvalues.

Corollary 1.3
Let a be a critical point of a C^2 function $f : \Omega \to \mathbb{R}$. If $\det H_f(a) \neq 0$, then f has either a local minimum or a local mximum or a saddle point at a.

Definition of degenerate critical points:
A critical point a of a C^2 function f is called non-degenerate if $\det H_f(a) \neq 0$ and degenerate otherwise.

Example of a degenerate critical point:
When $f(x, y) = x^3$ then $(0, 0)$ is a degenerate critical point of f, and f has neither a local extremum at $(0, 0)$ nor a saddle point.

Definition of the principal minors of a matrix:
Let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ matrix. Given $k = 1, 2, \ldots, n$, we will denote by A_k the $k \times k$ submatrix $A_k = (a_{ij})_{i,j=1}^k$.
The determinants $\det A_k$ are called the **principal minors of A**.

Proposition 1.4
Let A be a symmetric $n \times n$ matrix with $\det A \neq 0$. Then:

1. A is positive definite $\iff$ $\det A_k > 0 \forall k = 1, 2, \ldots, n$.
2. A is negative definite $\iff$ $(-1)^k \det A_k > 0 \forall k = 1, 2, \ldots, n$.
3. A is indefinite $\iff$ A is neither positive definite nor negative definite.

Corollary 1.5
Let $A = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$. Then:

1. A is positive definite $\iff$ $\alpha > 0$ and $\alpha\gamma - \beta^2 > 0$
2. A is negative definite $\iff$ $\alpha < 0$ and $\alpha\gamma - \beta^2 > 0$
3. A is indefinite $\iff$ $\alpha\gamma - \beta^2 < 0$

Example of classifying a critical point:
We found that the function $f(x, y) = xye^{-x^2-y^2}$ has 5 critical points: $\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right)$, $\left(\pm\frac{1}{\sqrt{2}}, \mp\frac{1}{\sqrt{2}}\right)$, and $(0, 0)$, with an absolute maximum at $\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right)$ and an absolute minimum at $\left(\pm\frac{1}{\sqrt{2}}, \mp\frac{1}{\sqrt{2}}\right)$.

Investigate the nature of $(0,0)$,

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\left[y(1-2x^2)e^{-x^2-y^2}\right] = -4xye^{-x^2-y^2} - 2xy(1-2x^2)e^{-x^2-y^2}$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left[x(1-2y^2)e^{-x^2-y^2}\right] = -4xye^{-x^2-y^2} - 2xy(1-2y^2)e^{-x^2-y^2}$$
$$\frac{\partial^2 f}{\partial x\partial y} = (1-2y^2)e^{-x^2-y^2} - 2x^2(1-2y^2)e^{-x^2-y^2}$$

So $H_f(0,0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is indefinite $\implies$ f has a saddle point at (0,0).

Example of non-degenerate critical points:
Find and classify the critical points of $f : \mathbb{R}^3 \to \mathbb{R}$ where $f(x,y,z) = x^3 - y^3 + 3xy + z^2 - 2z$.
$\nabla f = (3x^2 + 3y, -3y^2 + 3x, 2z - 2) = 0 \implies x^2 = -y, y^2 = x, z = 1 \implies x^2 = -y, x^4 = x, z = 1 \implies x^2 = -y, x = 0$ or $x = 1, z = 1 \implies (x,y,z) = (0,0,1), (1,-1,1)$.

So $(0,0,1)$ and $(1,-1,1)$ are the critical points. We have $H_f(x,y,z) = \begin{bmatrix} 6x & 3 & 0 \\ 3 & -6y & 0 \\ 0 & 0 & 2 \end{bmatrix}$,

so $H_f(0,0,1) = \begin{bmatrix} 0 & 3 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is clearly indefinite since the first principal minor is 0 and $H_f(1,-1,1) = \begin{bmatrix} 6 & 3 & 0 \\ 3 & 6 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is positive definite.

So we have non-degenerate critical points (as $\det H_f \neq 0$). Hence, $(0,0,1)$ is a saddle point; $(1,-1,1)$ is a local minimum.
But f has no global extrema because $f(x,0,0) = x^3$ can take arbitrarily positive and negative values.

Example of a degenerate critical point:
Let $f(x,y) = x^4 + y^4$ (with $(x,y) \in \mathbb{R}^2$).
$\nabla f = (4x^3, 4y^3) = 0 \iff (x,y) = (0,0)$.
$H_f(x,y) = \begin{bmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{bmatrix}, H_f(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
So $(0,0)$ is a degenerate critical point and the 2nd derivative test does not apply.
However, f has a global minimum at $(0,0)$.

Week 2

Inverse Function Theorem and Implicit Function Theorem

Theorem 2.1
Let $I \subseteq \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ is a continuous injective function. Then:

1. f is either strictly increasing or strictly decreasing.
2. $f(I)$ is an interval containing the same number of endpoints as I.
3. f is a homeomorphism of I onto $f(I)$.

Proof.

1. Let us first consider the case that $I = [a,b](a < b)$. Since f is injective, either $f(a) < f(b)$ or $f(b) < f(a)$. Assume that $f(a) < f(b)$ (the other case can be done symmetrically). Let's show that f is strictly increasing on $[a,b]$, i.e., $f(x) < f(y)$ whenever $a \leq x < y \leq b$. We argue by contradiction, supposing that $f(x) > f(y)$ for some $a \leq x < y \leq b$.
   Note that $f(y) > f(a)$, for otherwise $f(y) < f(a) < f(b)$ and by the Intermediate Value Theorem (IVT), $\exists \alpha \in (y,b)$ such that $f(\alpha) = f(a)$, contradicting the injectivity of f. Therefore $f(a) < f(y) < f(x)$ and so, again, by the IVT $\exists y' \in (a,x)$ such that $f(y') = f(y)$, again contradicting the injectivity of f.
   Next, let I be any interval. Pick up any $a,b \in I$ with $a < b$. Suppose that $f(a) < f(b)$ (the case $f(a) > f(b)$ can be done symmetrically). By the previous paragraph, we know that f is strictly increasing on $[a,b]$. Now, if $x,y \in I$ and $x < y$, then with $\alpha = \min\{a,x\}, \beta = \max\{y,b\}$, we have $[a,b],[x,y] \subseteq [\alpha,\beta] \subseteq I$. Since f is strictly increasing on $[a,b]$, we must have (using the 1st paragraph again) $f(\alpha) < f(\beta)$ and f is strictly increasing on $[\alpha,\beta]$. Hence, we conclude that f is strictly increasing on I.

2. Since f is continuous, $J = f(I)$ is an interval. Suppose that f is strictly increasing. Note that the inverse function f^{-1} is then also strictly increasing.
   Now, if I contains its left endpoint a, then $\forall x \in I$, $f(a) \leq f(x)$, so $f(a)$ is a left endpoint of J. Similarly, if I contains its right endpoint b, then $f(b)$ is the right endpoint of J. Applying the same argument with f^{-1} in place of f, we conclude if I contains its left (respectively, right) endpoint c, then $f^{-1}(c)$ is the left (respectively, right) endpoint of I. It follows that I and J contain the same number of endpoints.

3. If $I = [a,b]$, then f is a homeomorphism of I onto $f(I)$ because of our general result about continuous injective functions on compact sets.
   Otherwise, it follows that $f|[a,b]$ is a homeomorphism onto $f([a,b])$ for any $a,b \in I$ with $a \leq b$. This implies that $f^{-1} : f(I) \to I$ is continuous (at any $y \in f(I)$).
   Indeed, let $y \in f(I)$ and consider any sequence (y_n) in $f(I)$ with $y_n \to y$. Then the set $S = \{y\} \cup \{y_n : n \in N\}$ is compact, so it has both a smallest element $c = f(a)$ and a largest element $d = f(b)$. Assuming that f is strictly increasing we must have $a \leq b$, and $f([a,b]) = [c,d] \supseteq S$. Since $f|[a,b]$ is a homeomorphism onto $[c,d]$ (i.e., $(f|[a,b])^{-1} = f^{-1}|[c,d]$ is continuous), we obtain $f^{-1}(y_n) = (f^{-1}|[c,d])(y_n) \to (f^{-1}|[c,d])(y) = f^{-1}(y)$. It follows that f^{-1} is continuous at any $y \in f(I)$.

$\square$

Theorem 2.2
Let f be a bijection of a non-zero interval $I \subseteq \mathbb{R}$ onto an interval $J \subseteq \mathbb{R}$. If f is differentiable at $a \in I$, $f'(a) \neq 0$, and f^{-1} is continuous at $f(a)$ and $(f^{-1})'(f(a)) = \frac{1}{f'(a)}$

(Sketch). $\square$

Definition of a diffeomorphism:
Let f be a bijection of an open subset $U \subseteq \mathbb{R}^n$ onto an open subset $V \subseteq \mathbb{R}^n$. If both f and f^{-1} are differentiable (on U and V respectively), then f is called a **diffeomorphism** of U onto V. If both f and f^{-1} are C^k functions ($k = 1, 2, \ldots, \infty$), then f is called a **diffeomorphism of class** C^k.

Corollary 2.3
Let f be a differentiable homeomorphism of an open subset $U \subseteq \mathbb{R}$ onto an open subset $V \subseteq \mathbb{R}$. If $f'(a) \neq 0$ for all $a \in U$, then f is a diffeomorphism of U onto V. Moreover, if $f \in C^k(U)$, then f is a C^k diffeomorphism.

Proof. If $b = f(a) \in V$ (where $a \in U$), then there exists an open interval $I \subseteq U$ such that $a \in I$. Then $f(I)$ is another open interval and $f|I$ is a homeomorphism onto $f(I)$ (by the Inverse Function Theorem), and $f|I$ satisfies the assumptions of the above theorem. Hence, $(f|I)^{-1} = f^{-1}|f(I)$ is differentiable at b. But this means that f^{-1} is differentiable at b. Since $b \in V$ is artbitrary, f^{-1} is differentiable on V and so f is a diffeomorphism.
We also have $\left(f^{-1}\right)'(b) = \frac{1}{f^{-1}(a)} = \frac{1}{f'(f^{-1}(b))}$ for any $b = f(a) \in V$. Thus, $\left(f^{-1}\right)' = \frac{1}{f'} \circ f^{-1}$. That f^{-1} is C^k when f is C^k follows by induction on $k = 1, 2, \ldots$: When $k = 1$, then $\frac{1}{f'}$ is continuous (as $f \in C^1(U)$), and f^{-1} is continuous, so $\left(f^{-1}\right)' = \frac{1}{f'} \circ f^{-1}$ is continuous. Assuming that our claim is true for C^k functions, consider $f \in C^{k+1}(U)$. Then $f' \in C^k(U)$, and as $f \in C^k(U)$, $f^{-1} \in C^k(V)$ by induction. Hence, $\left(f^{-1}\right)' = \frac{1}{f'} \circ f^{-1}$ is a C^k function as the composition of two C^k functions. Therefore $f^{-1} \in C^k(V)$ $\square$

Corollary 2.4 Inverse Function Theorem in 1 variable
Let $I \subset \mathbb{R}$ be an open interval and $f : I \to \mathbb{R}$ a C^k function such that $f'(x) \neq 0$ for all $x \in I$. Then f is a C^k diffeomorphism of I onto $f(I)$.

Proof. By the IVT either $f'(x) > 0$ for all $x \in I$ (i.e., f is strictly increasing) or $f'(x) < 0$ for all $x \in I$ (i.e., f is strictly decreasing). Hence, f is injective and is a homeomorphism of I onto an open interval J. The assumption of the previous corollary are satisfied, hence the conclusion. □

Corollary 2.5 Inverse Function Theorem in 1 variable, local version
Let $U \in \mathbb{R}$ be open and $f : U \to \mathbb{R}$ be a C^k function. If $f'(a) \neq 0$ at some $a \in U$, then there exists an open interval I such that $a \in I \subseteq U$ and $f|I$ is a C^k diffeomorphism of I onto $f(I)$

How do these results generalize to functions of n variables?

Theorem 2.6
Let $\Omega \subseteq \mathbb{R}^n$ be an open set and let $f : \Omega \to \mathbb{R}^n$ be injective. Then $f(\Omega)$ is open and f is a homeomorphism of Ω onto $f(\Omega)$.

Proof. Omitted due to high difficulty. □

Lemma 2.7
If $T : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear transformation then there exists a $c > 0$ such that for all $x \in \mathbb{R}^n$, $\|T(x)\| \geq C\|x\|$

Proof. Recall that T^{-1} is a Lipschitz function, i.e., there exists $M > 0$ such that $\|T^{-1}(x)\| \leq M\|x\|$ for all $x \in \mathbb{R}^n$. Hence, for all $x \in \mathbb{R}^n$, $\|x\| = \|T^{-1}(T(x))\| \leq M\|T(x)\|$, so $\|T(x)\| \geq \frac{1}{M}\|x\|$. □

Theorem 2.8
Let f be a bijection of an open subset $U \subseteq R^n$ onto an open subset $V \in \mathbb{R}^n$. If f is differentiable at $a \in U$, $\det(D_f(a)) \neq 0$, and f^{-1} is continuous at $b = f(a)$, then f^{-1} is differentiable at b and $D_{f^{-1}}(b) = (D_f(a))^{-1}$.

Proof. Let $T = D_f(a)$, $b = f(a)$. It suffices to show that

$$\lim_{y\to b} \frac{f^{-1}(y) - f^{-1}(b) - T^{-1}(y-b)}{\|y-b\|} = 0$$

But,

$$\frac{f^{-1}(y) - f^{-1}(b) - T^{-1}(y-b)}{\|y-b\|} = -T^{-1}\left(\frac{y-b-T\left(f^{-1}(y)-f^{-1}(b)\right)}{\|y-b\|}\right)$$

So it suffices to show that

$$\lim_{y\to b} \frac{y-b-T\left(f^{-1}(y)-f^{-1}(b)\right)}{\|y-b\|} = 0$$

and this will be done if we show that

$$\lim_{k\to\infty} \frac{y_k-b-T\left(f^{-1}(y_k)-f^{-1}(b)\right)}{\|y_k-b\|} = 0$$

For every sequence $(y_k) \in V\setminus\{b\}$ with $y_k - b$. Let $x_k = f^{-1}(y_k) \in U\setminus\{a\}$ (i.e., $y_k = f(x_k)$). Then $x_k \to f^{-1}(b) = a$ because f^{-1} is continuous at b. Thus we need to show that

$$\lim_{k\to\infty} \frac{f(x_k)-f(a)-T(x_k-a)}{\|f(x_k)-f(a)\|} =$$

$$\lim_{k\to\infty}\left[\frac{\|x_k-x\|}{\|f(x_k)-f(a)\|}\frac{f(x_k)-f(a)-T(x_k-a)}{\|x_k-a\|}\right] = \lim_{k\to\infty} A_kB_k = 0$$

Now, as $T = D_f(a)$, $\lim_{k\to\infty} B_k = 0$ (by the definition of the derivative). So to complete the proof it is enough to show that the sequence (A_k) is bounded. But

$$\frac{1}{A_k} = \|\frac{f(x_k)-f(a)-T(x_k-a)}{\|x_k-a\|} + T\left(\frac{x_k-a}{\|x_k-a\|}\right)\| =$$

$$\|B_k + T\left(\frac{x_k-a}{\|x_k-a\|}\right)\| \geq \|T\left(\frac{x_k-a}{\|x_k-a\|}\right)\| - \|B_k\|$$

and by the lemma, there exists a $c > 0$ such that $\|T\left(\frac{x_k-a}{\|x_k-a\|}\right)\| \geq c$ for all k. As $B_k \to 0$, there exists a k_0 such that for all $k > k_0$ $\frac{1}{A_k} \geq \frac{c}{2}$ and so for all $k \in \mathbb{N}$ $\frac{1}{A_k} \geq \min\left\{\frac{c}{2}, \frac{1}{A_1}, \frac{1}{A_2}, \ldots, \frac{1}{A_{k_0}}\right\} > 0$. Hence, (A_k) is bounded. $\square$

Corollary 2.9
Let f be a differentiable homeomorphism of an open subset $U \subseteq \mathbb{R}^n$ onto an open subset $V \subseteq \mathbb{R}^n$. If $\det(D_f(x)) \neq 0$ for all $x \in U$, then f is a diffeomorphism of U onto V. Moreover, if $f \in C^k(U)$ then f is a C^k diffeomorphism.

Proof. Clearly, the assumptions of the previous theorem are satisfied for each $a \in U$, so f^{-1} is differentiable at each $b = f(a)$, and f is thus a diffeomorphism of U onto V. $\square$

Remark: The following example shows that the 1-dimensional Inverse Function Theorem cannot be generalized to n-dimensions.

Example of Polar Coordinate Mapping:
Let $f : (0, \infty) \times \mathbb{R}$ be given by $f(s, t)$

Theorem 2.10 Inverse Function Theorem (IFT)
Let $f : \Omega \to \mathbb{R}^n$ be a C^k function where $\Omega \subseteq \mathbb{R}^n$ is open (and $k = 1, 2, \ldots, \infty$). If $\det(D_f(a)) \neq 0$ for some $a \in \Omega$, then there exists an open set $U \in \Omega$ with $a \in U$ and an open set $V \subseteq \mathbb{R}^n$ with $f(a) \in V$ such that $f|U$ is a C^k diffeomorphism of U onto V.

Corollary 2.11 Open Mapping Theorem
Let $F : \Omega \to \mathbb{R}^n$ be C^1 function where $\Omega \subseteq \mathbb{R}^n$ is open. If $\det(D_f(x)) \neq 0$ for all $x \in \Omega$, then f is an open wrapping, i.e., for every open subset $W \subseteq \Omega$, $f(W)$ is open in $\mathbb{R}^n$.

Proof. Let $W \subseteq \Omega$ be open. To conclude that $f(W)$ is open, it suffices to show that for all $b \in f(W)$ there exists an open V such that $b \in V \subseteq f(W)$. But $b = f(a)$ for some $a \in W$ and $f|W$ and $a \in W$ satisfy the assumption of the IFT. Thus, there exists open $U \subseteq W$ and open $V \subseteq \mathbb{R}^n$ such that $a \in U$, $b \in V$ and $f(U) = (f|W)(U) = V$. Clearly, $b \in V \subseteq f(W)$. $\square$

Corollary 2.12
Let $f : \Omega \to \mathbb{R}^n$ bw a C^k function where $\Omega \to \mathbb{R}^n$ is open. If f is injective and $\det(D_f(x)) \neq 0$ for all $x \in \Omega$, then $f(\Omega)$ is open and f is a C^k diffeomorphism of Ω onto $f(\Omega)$.

Proof. By a previous corollary, it suffices to show that $f(\Omega)$ is open and f is a homeomorphism of Ω onto $f(\Omega)$. But by the previous corollary, f is an open mapping, so, in particular, $f(\Omega)$ is open. Thus, it remains to prove that $f^{-1} : f(\Omega) \to \Omega$ is continuous. Recall that this will be true if for each open $U \subseteq R^n$, $\left(f^{-1}\right)^{-1}(U)$ is open relative to $f(\Omega)$, i.e., is open in $\mathbb{R}^n$ because $f(\Omega)$ is open. But $\left(f^{-1}\right)^{-1}(U) = \left(f^{-1}\right)^{-1}(U \cap \Omega) = f(U \cap \Omega)$ is indeed open in R^n by the Open Mapping Theorem. $\square$

Example of determining a diffeomorphism:
The polar coordinate mapping $f(r,\theta) = (rcos\theta, rsin\theta)$ (considered on $(0,\infty) \times \mathbb{R}$), is an open mapping of $(0,\infty) \times \mathbb{R}$ onto $\mathbb{R}^2 \setminus \{(0,0)\}$ because $\det(D_f(r,\theta)) = r > 0$ for all $(r,\theta) \in (0,\infty) \times \mathbb{R}$.
Note that $\varphi = f|((0,\infty) \times (-\pi,\pi))$ is injective. Hence, by the last corollary φ is a C^∞ diffeomorphism on $(0,\infty) \times (-\pi,\pi)$ onto $\varphi((0,\infty) \times (-\pi,\pi)) = \mathbb{R}^2 \setminus ((-\infty,0] \times \mathbb{R})$.

$$D_{\varphi^{-1}}(rcos\theta, rsin\theta) = \begin{bmatrix} cos\theta & -rsin\theta \\ sin\theta & rcos\theta \end{bmatrix}^{-1} = \frac{1}{r}\begin{bmatrix} rcos\theta & rsin\theta \\ -sin\theta & cos\theta \end{bmatrix}$$

Similarly $\varphi|((0,\infty) \times (a,b))$, where $b - a = 2\pi$ is a c^∞ diffeomorphism on $(0,\infty) \times (a,b)$ onto $\mathbb{R}^2 \setminus \{r(cos\theta, sin\theta) : r \geq 0\}$.

Definition of an implicit function:
Let $\Omega_n \subseteq \mathbb{R}^n$, $\Omega_m \subseteq \mathbb{R}^m$, $F : \Omega_n \times \Omega_m \to \mathbb{R}^m$, and $c \in \mathbb{R}^m$.
Consider the equation

$$F(x,y) = c \qquad\qquad (x \in \Omega_n, y \in \Omega_m)(*)$$

which we suppose needs to solved for y. If for every $x \in \Omega_n$ this equation has a solution, then by choosing for each $x \in \Omega_n$ a solution $y \in \Omega_m$ and calling it $f(x)$, we obtain a function $f : \Omega_n \to \Omega_m$ such that $F(x.f(x)) = c$ for all $x \in \Omega_n$. Any such function is called an **implicit function** defined by Eq. $(*)$.

Note: If for all $x \in \Omega_n$ there exists a unique $y \in \Omega_m$ such that $F(x,y) = c$, then Eq. $(*)$ defines a unique implicit function, but in general, implicit functions are not unique.

Example of:
Let $n = m = 1$, $\Omega_n = \Omega_m = [-1,1]$, $F(x,y) = x^2 + y^2$, $c = 1$. Then the functions $f_\pm(x) = \pm\sqrt{1-x^2}$ are implicit functions defined by $(*)$ (i.e., eg. $x^2 + y^2 = 1$) and there are many other implicit functions.
If we replace Ω_m by $[0,1]$, then f_+ will be the unique implicit function defined by $(*)$ $(f_+(x) = \sqrt{1-x^2})$.

Question
Under what conditions does an implicit function exist; is unique; is it differentiable? If it is differentiable how can we obtain its derivative?

Note: Let $F : \Omega \to \mathbb{R}^m$ be a C^k function where $\Omega \subseteq \mathbb{R}n + m = \mathbb{R}^n \times R^m$ is open. We will write the elements of $\mathbb{R}^n + m = R^n \times R^m$ as (x, y) where $x \in R^n$, $y \in R^m$. Then

$$D_f(x,y) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1}(x,y) & \dots & \frac{\partial F_1}{\partial x_n}(x,y) & \frac{\partial F_1}{\partial y_1}(x,y) & \dots & \frac{\partial F_1}{\partial y_m}(x,y) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(x,y) & \dots & \frac{\partial F_m}{\partial x_n}(x,y) & \frac{\partial F_m}{\partial y_1}(x,y) & \dots & \frac{\partial F_m}{\partial y_m}(x,y) \end{bmatrix}$$

with the first $m \times n$ block will be named $\frac{\partial F}{\partial x}(x, y)$ and the second $m \times m$ block will be named $\frac{\partial F}{\partial y}(x, y)$.
Thus, we can write $D_F(x,y) = \begin{bmatrix} \frac{\partial F}{\partial x}(x,y) & \frac{\partial F}{\partial y}(x,y) \end{bmatrix}$

Theorem 2.13 Implicit Function Theorem (IPFT)
Let $F : \Omega \to \mathbb{R}^m$ be a C^k function where $\Omega \subseteq \mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ is open. Suppose that for $(a, b) \in \Omega$ and $c \in \mathbb{R}^m$, $F(a, b) = c$ and $\det\left(\frac{\partial F}{\partial y}(a, b)\right) \neq 0$. Then there exist open sets $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ that satisfy:

1. $(a, b) \in U \times V$,
2. for all $x \in U$, there exists a unique $y \in V$ such that $F(x, y) = c$.

Moreover, the unique implicit function $f : U \to V$ defined by the equation $F(x, y) = c$ ($x \in U$, $y \in V$) is a C^k funciton.

Proof. Define $G : \Omega \to \mathbb{R}^{n+m}$ by $G(x,y) = (x, F(x.y))$. This is a C^k function, $G(a,b) = (a,c)$ and

$$D_G(x,y) = \begin{bmatrix} I_n & 0 \\ \frac{\partial F}{\partial x}(x,y) & \frac{\partial F}{\partial x}(x,y) \end{bmatrix}$$

Thus $\det(D_G(a,b)) = (\det I_n)\left(\det\left(\frac{\partial F}{\partial y}(a,b)\right)\right) \neq 0$.
Thus by the IFT, there exists an open subset $\Omega_1 \subseteq \Omega$ with $(a,b) \in \Omega_1$ and an open subset $\Omega \subseteq \mathbb{R}^{n+m}$ with $(a,c) = G(a,b) \in W$ such that $G|\Omega_1$ is a C^k diffeomorphism of Ω_1 onto W. Let $H = (G|\Omega_1)^{-1} : W \to \Omega_1$. Then $H(x,y) = (j(x,y), k(x,y))$ where $j : W \to \mathbb{R}^n$ and $k : W \to \mathbb{R}^m$ are C^k functions. Note that $(x,y) = G(H(x,y)) = (j(x,y), F(k(x,y)))$ for all $(x,y) \in W$. Hence, $j(x,y) = x$ and $F(k(x,y)) = y$ for all $(x,y) \in W$. Thus $H(x,y) = (x, k(x,y))$ and so for all $(x,y) \in W$,

$$(x, k(x,y)) \in \Omega_1 \text{ and } F(x, k(x,y)) = y$$

Note that we may assume that $\Omega_1 = U' \times V$ where $U' \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ are open. [Indeed, $(a,b) \in \Omega_1$ and Ω_1 is open, so there exists an $r > 0$ such that $B_r^{n+m}(a,b) \in \Omega_1$. But $B_r^{n+m}(a,b) \supseteq B^n_{\frac{r}{2}}(a) \times B^m_{\frac{r}{2}}(b)$. So we can take $U' = B^n_{\frac{r}{2}}(a)$, $V = B^m_{\frac{r}{2}}(b)$ and replace Ω_1 with $U' \times V$ and W with $G(U' \times V)$].
Moreover, since $(a,c) \in W$, we can find an open set U such that $a \in U \subseteq U'$ and $U \times \{c\} \subseteq W$. Then for all $x \in U$, $(x,c) \in W$ and so $F(x, k(x,c)) = c$. Thus when $f : U \to V$ is given by $f(x) = k(x,c)$, then f is an implicit funciton defined by the equation $F(x,y) = c$ (for $x \in U$, $y \in V$). It is clear that f is a C^k function.
It remains to confirm that for all $x \in U$ there exists a unique $y \in V$ such that $F(x,y) = c$. But if $y_1, y_2 \in V$ and $F(x,y_1) = c = f(x,y_2)$, then $G(x,y_1) = (x,c) = G(x,y_2)$, and so $y_1 = y_2$ as $G|U \times V$ is injective. $\square$

Week 3

(no title yet)

Corollary 3.1
With the assumptions and notation of the IPFT, let $S = \{(x,y) \in \Omega : F(x,y) = c\}$. Then $S \cap (U \times V) = \{(x,y) \in \mathbb{R}^{n+m} : x \in U \text{ and } y = f(x)\}$.

Remark: Note that when $m = 1$, then $\det\left(\frac{\partial F}{\partial y}\right) = \frac{\partial F}{\partial y}$. So if $\frac{\partial F}{\partial y}(a,b) \neq 0$ then the level set $S = \left\{(x,y) \in \mathbb{R}^{n+1} : F(x,y) = c\right\}$ in a neighbourhood of (a,b) is the graph of the implicit function.

Example of:
(IPFT, level set, and graph) Consider the level set $S = \left\{(x,y) \in \mathbb{R}^2 : x^3y^2 + y^3(x-1)^2 = 1\right\}$ of $F(x,y) = x^3y^2 + y^3(x-1)^2$.

1. Show that S is not the graph of any function $y = f(x)$, i.e., $S \neq \left\{(x,y) \in \mathbb{R}^2 : y = f(x)\right\}$.

2. Show that in a neighbourhood of $(1,1)$, S is the graph of a smooth function f and find the slope of the tangent line to the graph of f at $(1,1)$.

Solutions:

1. $(1,-1), (1,1) \in S$, so no such function exists.

2. $\frac{\partial F}{\partial y}(1,1) = 2x^3y + 3^2(x-1)^2\Big|_{x=1,y=1} = 2 \neq 0$. So by the IPFT (with $a = b = c = 1$) and the corollary there exist open sets $U, V \in \mathbb{R}$ with $(1,1) \in U \times V$ and a smooth function $f : U \to V$ such that $f(1) = 1$, $F(x, f(x)) = 1 = x^3 f(x)^2 + f(x)^3 (x-1)^2 = 1$ for all $x \in U$, and $S \cap (U \times V) = \{(x,y) : x \in U \text{ and } y = f(x)\}$.

   The slope is $f^{-1}(1)$: Since $x^3 f(x)^2 + f(x)^3 (x-1)^2 = 1$ for all $x \in U$, so $0 = \frac{\mathrm{d}}{\mathrm{d}x}\left[x^3 f(x)^2 + f(x)^3 (x-1)^2\right] = 3x^2 f(x)^2 + 2x^3 f(x) f'(x) +$

$3f(x)^2 f'(x)(x-1)^2 + 2f(x)^3(x-1)$. When $x=1$, $f(1)=1$, and so $0 = 3 + 2f'(1)$. Thus $f'(1) = \frac{3}{2}$

Example of:
(Finding the derivative without the function) Consider the problem of solving the system of equations: $\begin{cases} xy^2 + xzu + yv^2 = 3 \\ u^3yz + 2xv - u^2v^2 = 2 \end{cases}$ (*). for u and v in terms of x, y, z near $x = y = z = u, v = 1$ and computing the partial $\frac{\partial u}{\partial z}, \frac{\partial v}{\partial z}$.

Let $a = (1,1,1), b = (1,1), c = (3,2)$, and $F : \mathbb{R}^3 \to \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$F(x,y,z,u,v) = (xy^2 + xzu + yu^2, u^3yz + 2xv - u^2v^2).$$

Then $F(a,b) = c$, $\frac{\partial F}{\partial(u,v)} = \begin{bmatrix} xz & 2yv \\ 3u^2yz - 2uv^2 & 2x - 2u^2v \end{bmatrix}$.

$$\det\left(\frac{\partial F}{\partial(u,v)}(a,b)\right) = \det\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} = -2 \neq 0.$$

Hence, by the IPFT there exists a smooth function $f(x,y,z) = (f_1(x,y,z), f_2(x,y,z))$ defined on a neighbourhood U of $u = (1,1,1)$ such that $F(x,y,z,f(x,y,z)) = (3,2) = c$ for all $(x,y,z) \in U$ and $f(1,1,1) = (1,1)$: $u = f_1(x,y,z), v = f_2(x,y,z)$ are the expressions of u and v in terms of x, y, z. To find $\frac{\partial u}{\partial z}$ and $\frac{\partial v}{\partial z}$ we differentiate Eqs(*) with respect to z, treating u and v as functions of x, y, z:

$$0 = \frac{\partial}{\partial z}\left(xy^2 + xzu + yv^2\right) = xu + xz\frac{\partial u}{\partial z} + 2yv\frac{\partial v}{\partial z}$$
$$0 = \frac{\partial}{\partial z}\left(u^3yz + 2xv - u^2v^2\right) = 3u^2\frac{\partial u}{\partial z}yz + u^3y + 2x\frac{\partial v}{\partial z} - 2u\frac{\partial u}{\partial z}v^2 - u^2 2v\frac{\partial v}{\partial z}$$

With $(x,y,z) = (1,1,1)$, $(u,v) = (1,1)$ we get

$$1 + \frac{\partial u}{\partial z} + 2\frac{\partial v}{\partial z} = 0, \frac{\partial u}{\partial z} + 1 = 0.$$

Hence, $\frac{\partial f_1}{\partial z} = \frac{\partial u}{\partial z}(1,1,1) = -1$, $\frac{\partial f_2}{\partial z} = \frac{\partial v}{partial z}(1,1,1) = 0$.

Proposition 3.2 Implicit Differentiation
Let $F : \Omega_n \times \Omega_m \to \mathbb{R}^m$ be a C^1 function where $\Omega_n \subset \mathbb{R}^n$ and $\Omega_m \subset \mathbb{R}^m$ are open and let $c \in \mathbb{R}^m$. If $f : \Omega_n \to \Omega_m$ is a differentiable function such that $F(x, f(x)) = c$ for all $x \in \Omega_n$, then

$$\frac{\partial F}{\partial y}(x, f(x)) D_f(x) = -\frac{\partial F}{\partial x}(x, f(x))$$

and

$$D_f(x) = -\left[\frac{\partial F}{\partial y}(x, f(x))\right]^{-1} \frac{\partial F}{\partial x}(x, f(x))$$

provided $\det\left(\frac{\partial F}{\partial y}(x, f(x))\right) \neq 0$.

Proof. Define $g : \Omega_n \to \Omega_n \times \Omega_m$ by $g(x, f(x))$. Then g is differentiable and

$$D_g(x) = \begin{bmatrix} I_n \\ D_f(x) \end{bmatrix}.$$

Since $(F \circ g)(x) = c$, the chain rule yields $0 = D_{F \circ g}(x) = D_F(g(x))\, D_g(x) = \begin{bmatrix} \frac{\partial F}{\partial x}(g(x)) & \frac{\partial F}{\partial y}(g(x)) \end{bmatrix} \begin{bmatrix} I_n \\ D_f(x) \end{bmatrix} = \frac{\partial F}{\partial x}(x, f(x)) + \frac{\partial F}{\partial y}(x, f(x))\, D_f(x)$.
Hence, the result. $\square$