Lecture 20:

Scaling marks.

Recall: Defn: $\bar{f} : \mathbb{R}^n \to \mathbb{R}^m$ is $C^1$ at $\bar{x}_0$ if each entry of $D\bar{f}$ exists in a nbhd $\bar{x}_0$ and is cts at $\bar{x}_0$.

Recall: DIAGRAM of implications Theorem:

\[
C^1 \Rightarrow \text{diffble} \Rightarrow \text{cts} \quad \downarrow \\
D\bar{f} \text{ exists}
\]

Recall that

\[
f(x, y) = \frac{xy}{x^2 + y^2}, \quad (x, y) \neq (0, 0)
\]

and $f(0, 0) = 0$ has partial derivatives everywhere but is not even continuous and therefore not diffble.

In fact, in HW6, you will see a function $f : \mathbb{R}^2 \to \mathbb{R}$ that has directional derivatives at every point and in every direction but is not even continuous and therefore not diffble.

Theorem 4 (p, 714): $C^1 \Rightarrow$ diffble

Proof: (in special case $f : \mathbb{R}^2 \to \mathbb{R}$)

Lemma (Theorem 3 (p, 714)): If $f_x$ and $f_y$ exist in a nbhd. of $(x_0, y_0)$ and $h, k$ are suff. small, then there exist $0 \leq \theta_1, \theta_2 \leq 1$ s.t.

\[
f(x_0 + h, y_0 + k) - f(x_0, y_0) = h f_x(x_0 + \theta_1 h, y_0 + k) + k f_y(x_0, y_0 + \theta_2 k)
\]

Proof: Write

\[
f(x_0 + h, y_0 + k) - f(x_0, y_0) = (f(x_0 + h, y_0 + k) - f(x_0, y_0 + k))
\]

\[+(f(x_0, y_0 + k) - f(x_0, y_0))
\]

For the second term apply the single variable MVT to the function $g(t) = f(x_0, y_0 + tk)$, $0 \leq t \leq 1$.

\[
g'(\theta_2) = f(x_0, y_0 + \theta_1 k) - f(x_0, y_0)
\]

\[K f_y(y_0 + \theta_2 k)
\]
For the first term apply the single variable MVT to the function \( f(x_0 + th, y_0 + k), \ 0 \leq t \leq 1. \)

\[
\frac{f(x_0 + th, y_0 + k) - f(x_0, y_0) - h f_x(x_0, y_0) + k f_y(x_0, y_0)}{\sqrt{h^2 + k^2}}
\]

\[
= \left| \frac{h}{\sqrt{h^2 + k^2}}(f_x(x_0 + \theta_1 h, y_0 + k) - f_x(x_0, y_0)) \right| \\
+ \left| \frac{k}{\sqrt{h^2 + k^2}}(f_y(x_0, y_0 + \theta_2 k) - f_y(x_0, y_0)) \right|
\]

which tends to 0 as \((h, k) \to (0, 0)\) because of continuity of \( f_x \) and \( f_y \) at \((x_0, y_0)\).

Theorem is important because it gives an easy way to verify that a function is diffble. and also because some results require \( C^1 \) instead of diffble., i.e., they are false with only the diffble assumption.

**Defn:** \( \overline{f} \) is \((C^1, \text{ diffble, cts})\) on a nbhd. \( I \) iff it is \((C^1, \text{ diffble, cts})\) at every point of the nbhd.

**Implicit functions:**

Consider an equation of the form \( F(x, y) = 0 \) where \( F \) is differentiable. If there is a differentiable function \( y = f(x) \) (an *implicit function*) which satisfies

\[
F(x, f(x)) = 0
\]
then by chain rule, we have
\[ \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} f'(x) = 0. \]

If also \( \frac{\partial F}{\partial y} \neq 0 \), then and thus
\[ f'(x) = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \]

Example: \( F(x, y) = y^2 - x \), then \( y = f(x) = \sqrt{x} \) is an implicit function. Indeed:
\[ -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = 1/(2y) = 1/(2\sqrt{x}) = f'(x) \]

\[ (y, y_0) \]

Implicit function theorem guarantees the existence of an implicit function, but it requires \( C^1 \) (Dibble. is not enough; you really need \( C^1 \)).

**One-dimensional Implicit Function Theorem:**
Let \( F(x, y) \) be a \( C^1 \) function on a nbhd. of \( (x_0, y_0) \) such that
\[ F(x_0, y_0) = 0 \]

and
\[ \frac{\partial F}{\partial y}(x_0, y_0) \neq 0 \]
then there is a function \( y = f(x) \) defined in a nbhd. \( I \) of \( x_0 \) such that on \( I \)
1. \( f \) is \( C^1 \)
2. \( f(x_0) = y_0 \),
3. on \( I \),
\[ f'(x) = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \]
4. \( F(x, f(x)) = 0 \) for \( x \in I \).

\[
\frac{\partial F}{\partial y} \neq 0 \iff \nabla F \text{ not horizontal} \iff \text{tangent to level curve not vertical}
\]

Higher-dimensional implicit function theorem is a generalization of \( f + 1 \).
Lecture 21:
Suppose we want the system of two equations in five unknowns
\[ F(x, y, z, u, v) = 0 \]
\[ G(x, y, z, u, v) = 0 \]
to determine \( u \) and \( v \) as functions of the independent variables \( x, y, z \):
\[ u = u(x, y, z), \quad v = v(x, y, z), \]
and we want to find all the partial derivatives \( u_x, u_y, u_z, v_x, v_y, v_z \).
Write
\[ R(x, y, z) = F(x, y, z, u(x, y, z), v(x, y, z)) = 0 \]
\[ S(x, y, z) = G(x, y, z, u(x, y, z), v(x, y, z)) = 0 \]

\( F_x, F_y, ..., G_x, G_y, ... \) mean the partial derivatives with respect to
the variables indicated and have nothing to do with the equations \( F = 0, G = 0 \).
\[ R(x, y, z) = 0 \] and so \( R_x = 0 \).

Use Chain Rule:
\[ F_x x_x + F_y y_x + F_z z_x + F_u u_x + F_v v_x = R_x = 0 \]
\[ G_x x_x + G_y y_x + G_z z_x + G_u u_x + G_v v_x = S_x = 0 \]
But since \( x_x = 1 \) and \( y_x = 0 = z_x \), this simplifies to:
\[ F_x + F_u u_x + F_v v_x = 0 \]
\[ G_x + G_u u_x + G_v v_x = 0 \]

View these as two linear equations with two unknowns: \( u_x \) and \( v_x \).

Eliminate \( v_x \):

--- Multiply 1st eqn by \( G_v \) and 2nd eqn. by \( F_v \) and then subtract equations and solve for \( u_x \):

\[
G_v F_x + G_v F_u u_x + G_v F_v v_x = 0
\]

\[
F_v G_x + F_v G_u u_x + F_v G_v v_x = 0
\]

\[
u_x = -\frac{F_x G_v - F_v G_x}{F_u G_v - F_v G_u}\]

The numerator and denominator can be viewed as determinants:

\[
u_x = -\frac{\frac{\partial(F,G)}{\partial(x,v)}}{\frac{\partial(F,G)}{\partial(u,v)}}\]

where

\[
\frac{\partial(F,G)}{\partial(x,v)} = \begin{vmatrix} F_x & F_v \\ G_x & G_v \end{vmatrix}
\]

\[
\frac{\partial(F,G)}{\partial(u,v)} = \begin{vmatrix} F_u & F_v \\ G_u & G_v \end{vmatrix}
\]

Denominator is determinant of partials of \( F, G \) w.r.t. dependent variables.

Numerator is the same except that you replace the appropriate dependent variable with the appropriate independent variable.

These are called Jacobian determinants.

Need: denominator to be \( \neq 0 \).

Get similar expressions for \( u_y, u_z, v_x, \) etc..
Example:

\[ x = u^2 + uv - v^2 \]
\[ y = 2uv + v^2 \]

Assume these eqns determine \( u = u(x, y), v = v(x, y) \). Find partial derivatives, say \( \frac{\partial u}{\partial x} \). Take partials of both eqns w.r.t. \( x \)

\[
1 = \frac{\partial x}{\partial x} = (2u + v) \frac{\partial u}{\partial x} + (u - 2v) \frac{\partial v}{\partial x}
\]
\[
0 = \frac{\partial y}{\partial x} = (2v) \frac{\partial u}{\partial x} + (2u + 2v) \frac{\partial v}{\partial x}
\]

View these as a system of two equations with two unknowns. Solve for \( \frac{\partial u}{\partial x} \) and \( \frac{\partial u}{\partial y} \). Mult. first eqn. by \( (2u + 2v) \) and second eqn. by \( (u - 2v) \), subtract, thereby eliminating \( \frac{\partial v}{\partial x} \), and then solve for \( \frac{\partial u}{\partial x} \).

\[
\frac{\partial u}{\partial x} = \frac{2u + 2v}{(2u + 2v)((2u + v) - 2v(u - 2v))}
\]

Theorem 8 (p. 740): Higher dimensional implicit function theorem.

Let \( F_1, \ldots, F_n \) be \( n \) functions of \( m + n \) variables:

\[ F_i(x_1, \ldots, x_m, y_1, \ldots, y_n) \]

\( n \) dependent variables \( y_1, \ldots, y_n \) and
\( m \) independent variables \( x_1, \ldots, x_m \).

Assume the \( F_i \) are \( C^1 \) in a nbhd. of a point \( \bar{x}_0 := (x_1^0, \ldots, x_m^0, y_1^0, \ldots, y_n^0) \).

Assume

\[
\frac{\partial(F_1, \ldots, F_n)}{\partial(y_1, \ldots, y_n)}(\bar{x}_0) \neq 0.
\]

Then there exist functions \( y_i = \phi_i(x_1, \ldots, x_m), i = 1, \ldots, n \) define on a nbhd \( I \) of \( (x_1^0, \ldots, x_m^0) \) s.t.
1. each $\phi_i$ is $C^1$ on $I$
2. $\phi_i(x_1^0, \ldots, x_m^0) = y_i^0$, 
3. on $I$, 
   \[
   \frac{\partial y_i}{\partial x_j} = - \frac{\frac{\partial (F_1, \ldots, F_i, \ldots, F_n)}{\partial (y_1, \ldots, x_j, \ldots, y_n)}}{\frac{\partial (F_1, \ldots, F_i, \ldots, F_n)}{\partial (y_1, \ldots, y_i, \ldots, y_n)}}
   \]
4. $F(x_1, \ldots, x_m, \phi_1((x_1, \ldots, x_m)), \ldots, \phi_n((x_1, \ldots, x_m)) = 0$ for all $x \in I$.

Inverse Function Theorem (2×2 version): Let $\overline{A}(z, w) = (f(z, w), g(z, w)) : \mathbb{R}^2 \to \mathbb{R}^2$ be a $C^1$ function on some nbhd. of some point $(z_0, w_0)$ and assume that 

\[
\frac{\partial (f, g)}{\partial (z, w)}(z_0, w_0) \neq 0.
\]

Then there is a function $\overline{B}(x, y) = (h(x, y), k(x, y)) : \mathbb{R}^2 \to \mathbb{R}^2$ defined in a nbhd. $N$ of $(x_0, y_0) = (f(z_0, w_0), g(z_0, w_0))$ such that

1. $\overline{B}$ is $C^1$ on $N$
2. $\overline{B}(x_0, y_0) = (z_0, w_0)$
3. $\frac{\partial (h, k)}{\partial (x, y)} = 1/\frac{\partial (f, g)}{\partial (z, w)}$ on $N$
4. $\overline{A} \circ \overline{B} = $ Identity on $N$
Lecture 22:
Re-state Inverse Function Theorem.

Proof: Apply the implicit function theorem to

\[ F(x, y, z, w) := x - f(z, w), \quad G(x, y, z, w) := y - g(z, w); \]

Check hypotheses for Implicit function theorem:
Both \( F \) and \( G \) are \( C^1 \) and

\[ \frac{\partial(F, G)}{\partial(z, w)} = \frac{\partial(f, g)}{\partial(z, w)}(z_0, w_0) \neq 0. \]

By implicit function theorem, we can implicitly solve for \( z = h(x, y), w = k(x, y), C^1 \) functions, to satisfy:

\[ F(x, y, h(x, y), k(x, y)) = 0, \quad G(x, y, h(x, y), k(x, y)) = 0 \]

which means means that

\[ x = f(h(x, y), k(x, y)), \quad y = g(h(x, y), k(x, y)). \]

And so \( \overline{B}(x, y) := (h(x, y), k(x, y)) \) is an inverse function to \( \overline{A}(z, w) := (f(z, w), g(z, w)) \).

\[ \mathbb{R}^2 \xrightarrow{\overline{B}} \mathbb{R}^2 \xrightarrow{\overline{A}} \mathbb{R}^2 \]

\[ x, y \quad \quad z, w \quad \quad x, y \]

Noting that

\[ D\overline{A} = \frac{\partial(f, g)}{\partial(z, w)}, \quad D\overline{B} = \frac{\partial(h, k)}{\partial(x, y)}, \]

we see that part 3 is follows from chain rule:

\[ 1 = |D(\text{Identity})| = |D(\overline{A} \circ \overline{B})| = |D(\overline{A})D(\overline{B})| = |D(\overline{A})||D(\overline{B})| \]
Main Example: Polar coordinates. \( \overline{A}(r, \theta) = (r \cos(\theta), r \sin(\theta)) \).

\[
x = f(r, \theta) = r \cos \theta, \quad y = g(r, \theta) = r \sin \theta
\]

In HW7.

Max and Min

Let \( f : D \subseteq \mathbb{R}^n \to \mathbb{R} \) and \( \overline{x}_0 \in D \).

(think \( n = 2 \))

Defn: \( f \) has a maximum (or global maximum) at \( \overline{x}_0 \) if \( \forall \overline{x} \in D \), \( f(\overline{x}_0) \geq f(\overline{x}) \). (similarly minimum or local minimum).

Defn: extremum : max or min.

Recall: \( B_r(\overline{x}_0) \) is the open ball centered at \( \overline{x}_0 \) with radius \( r \).

Defn: \( f \) has a local maximum (or relative maximum) at \( \overline{x}_0 \) if \( \exists r > 0 \) s.t. \( \forall \overline{x} \in B_r(\overline{x}_0) \cap D \), \( f(\overline{x}_0) \geq f(\overline{x}) \). (similarly local minimum).

Defn: local extremum : local max or local min.

Note: any global extremum is a local extremum.

Terminology: extremum is a point in the domain; extreme value (or just value) is the value of the function at the extremum.

Example of local max that is not a global max: \( f : \mathbb{R} \to \mathbb{R} \).
— a cubic polynomial has neither a global max nor a global min; but it can have two turning points that are local max and local min.

— a 4th degree polynomial has a global max or a global min but not both.

— a 4th degree polynomial can have three turning points, including one global max, one local max and one local min.

Theorem 1 (p. 753) Necessary conditions for local extrema:
Any local extremum must be one of the following types:

a. critical point: $\nabla f(\vec{x}_0) = \vec{0}$.

or

b. singular point: $\nabla f(\vec{x}_0)$ does not exist (i.e., one of the partials does not exist at $\vec{x}_0$).

or

c. boundary point: $\vec{x}_0 \in \partial D$.

Recall

$$\partial D = \{\vec{x}_0 \text{ s. t. for all } r > 0, B_r(\vec{x}_0) \cap D \neq \emptyset \text{ and } B_r(\vec{x}_0) \cap D^c \neq \emptyset \}.\}$$

Proof:

Let $\vec{x}_0 \in D$ be a local extremum.

If $\vec{x}_0$ is not a boundary point, then there exists $r > 0$ s.t. $B_r(\vec{x}_0) \subseteq D$.

If $\vec{x}_0$ is not a singular point, then $\nabla f(\vec{x}_0)$ must exist.

We must show that $\nabla f(\vec{x}_0) = \vec{0}$.

Proof by contradiction: Suppose $\nabla f(\vec{x}_0) \neq \vec{0}$.

For all sufficiently small $t > 0$, $\vec{x} := \vec{x}_0 + t\nabla f(\vec{x}_0) \in B_r(\vec{x}_0) \subseteq D$.

Since $\nabla f(\vec{x}_0)$ points in direction of increasing values of $f$, for $t > 0$,

$$f(\vec{x}) > f(\vec{x}_0)$$

and for $t < 0$,

$$f(\vec{x}) < f(\vec{x}_0).$$

Thus, $f$ cannot have a local extremum at $\vec{x}_0$. Thus, $\nabla f(\vec{x}_0) = \vec{0}$, as desired.