Last week, we explored local extrema for multivariable functions (or on a manifold):

If $f$ is in $C^2(U)$, $U \subseteq \mathbb{R}^n \to \mathbb{R}$, then find points $a \in U$ s.t. $[Df(a)] = 1 \times n$ 0-matrix. 
Examine quadratic form coming from $P^2$ f.s.

This week: Optimize the function when constrained to submanifold $M \subseteq \mathbb{R}^n$.

In pictures: $f: \mathbb{R}^2 \to \mathbb{R}$. $M$: unit circle, 1-dim'l manifold given by $x^2 + y^2 - 1 = 0$ in $\mathbb{R}^2$.

Problem: Max along $M$ may not be local max on graph of $f(x,y)$.

But, knowing it is a local extremum on $M$ means that along vectors tangent to $M$,

it is true that $[Dv(a)] = 0$.

Rephrased: "derivative vanishes on tangent space to manifold" at critical point $a$.

**Theorem:** $M \subseteq \mathbb{R}^n$ manifold, $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ constrained is $C^1$ function. $e \in M \cap U$ is a local extremum of $f|_M$.

then $T_e M \subseteq \ker [Df(e)]$. 

\[
\begin{align*}
\text{Diagram:} & \\
\text{Sketch of graph of $f(x,y)$ with unit circle $M$.} & \\
\text{Sketch of tangent space to $M$ at critical point $a$.} & \\
\end{align*}
\]
$T_e M = \text{tangent space to manifold}$ \text{ def } \Rightarrow \text{ "slope" in tangent line, in that its graph is a subspace of } \mathbb{R}^n \text{ viewed as linear trans.}

Main result in §3.2 was elegant characterization of $T_e M$ as

$\ker \left[ DF(e^1) \right]$ \text{ where } $M$ is

zero locus of $F$. (Don't be confused here:

$F(e) = 0$ defines manifold, locally written as

explicit function $\phi$.

So equivalently, the theorem says $\ker \left[ DF(e^1) \right] \subseteq \ker \left[ DF(e^2) \right]$ at constrained critical point.

\text{if } $M$ is a $2$-manifold

Do first example in book (3.7.3)

$f(x,y) = xy$ on unit circle with $x, y > 0$. (i.e. in first quadrant)

open quarter of circle is manifold.

as is any open set of manifold.

compute two kernels:

$F = x^2 + y^2 - 1$

$DF (x_0, y_0) = \begin{bmatrix} 2x_0 & 2y_0 \end{bmatrix}$\hspace{0.5cm} $\ker \begin{bmatrix} 2x_0 & 2y_0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$ \text{ such that } $2x_0 a + 2y_0 b = 0$.

$Df (x_0 y_0) = \begin{bmatrix} y_0 & x_0 \end{bmatrix}$\hspace{0.5cm} $\ker \begin{bmatrix} y_0 & x_0 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$ \text{ such that } $ay_0 + bx_0 = 0$.

Note: In this case, only possibility is $\ker (DF(e^1)) = \ker (DF(e^2))$ when are they the same?
We could have predicted this by thinking about when graph \( f(x,y) = xy = m \).

\( xy = m \) is level curve in \( xy \)-plane on which \( f(x,y) = m \).

Further we go from origin, larger fraction becomes.

When \( xy = m \) is tangent to unit circle, this should be max.

(No min on open quarter of circle, since mins occur at points in closure \( x = 0 \) or \( y = 0 \)).

When does \( xy = m \) intersect \( x^2 + y^2 = 1 \) in single point?

Substitute: \( x^2 + \left( \frac{m}{x} \right)^2 = 1 \)

Solve: \( x = \frac{\sqrt{2}}{2} = \cos \frac{\pi}{4} \)

Then \( y \), on unit circle, is also \( \frac{\sqrt{2}}{2} = \sin \frac{\pi}{4} \).

I.e. \( x_0 = y_0 \) as predicted above.

Criterion seems good start, but seems it may be hard to figure out what \( \epsilon \) to choose to get \( \ker \{ DF(\epsilon) \} \subset \ker \{ DF(\epsilon) \} \)

(this is where Lagrange multipliers come in. Give us a checkable condition for when this occurs)

pf of theorem: \( M \) is locally representable as \( \begin{bmatrix} g(y) \\ y \end{bmatrix} \) with \( y \) indep. vars in \( \mathbb{R}^k \)

Write \( \epsilon = \begin{bmatrix} a & b \end{bmatrix} \) as usual with \( a = g(b) \).

Then if \( \epsilon \) is constrained critical point, we have \( D(f \circ \tilde{g})(b) = 0 \).

\[ g(\tilde{y}) : \mathbb{R}^k \to \mathbb{R}^n \]

Points of \( M \) are image(\( \tilde{g} \)).
In other words, \( b \) is honest local extremum for composition \( f \circ \tilde{g} \).

On other hand, we compute \( D (f \circ \tilde{g}) \) by chain rule:

\[
\begin{align*}
= \begin{bmatrix} Df (\tilde{g}(b)) \end{bmatrix} \begin{bmatrix} D\tilde{g} (b) \end{bmatrix} & \text{so if this } = 0, \\
\ker \begin{bmatrix} Df (c) \end{bmatrix} \supseteq \text{im} \begin{bmatrix} D\tilde{g} (b) \end{bmatrix} = \text{graph of } \begin{bmatrix} Dg (b) \end{bmatrix} \\
= TcM. & \\
\end{align*}
\]

Problem yet to be solved: Find \( \epsilon \in U \cap M \)

such that \( \ker \begin{bmatrix} Df (c) \end{bmatrix} \supseteq \ker \begin{bmatrix} D\tilde{g} (c) \end{bmatrix} \).

\( f : U \subseteq \mathbb{R}^n \to \mathbb{R} \)

function we're optimizing

so \( Df (c) \) is \( 1 \times n \)

\( \text{smooth } \eta \)

\( \text{defined in } \mathbb{R}^n \)

\( \text{open set } U \)

As linear transformations, when is \( \ker \begin{bmatrix} B \end{bmatrix} = \ker \begin{bmatrix} A \end{bmatrix} \).

\( B : 1 \times n \to \mathbb{R} \)

\( A : m \times n \to \mathbb{R} \)

Think of \( A \) is \( m \times 1 \) matrices \( x_1, \ldots, x_m \).

Claim: \( \ker \begin{bmatrix} B \end{bmatrix} = \ker \begin{bmatrix} A \end{bmatrix} \iff \exists \lambda_1, \ldots, \lambda_m \in \mathbb{R} \text{ s.t. } \\
B = \lambda_1 x_1 + \cdots + \lambda_m x_m \)

\((\Leftarrow) \text{ is immediate} \)

\((\Rightarrow) \text{ next time...} \)