Take Home Midterm Exam MATH 939, Due Nov 14

1. Let $f : I \to I$ be a Markovian map on $I = [0, 1]$ with partition $I = \bigcup_{i=1}^n I_i$ with $n \geq 2$. Prove the following identities.

a. $|f^{-1}(I_k) \cap I_i|\,|f^{-1}(I_j) \cap I_k| = |f^{-1}(f^{-1}(I_j) \cap I_k) \cap I_i|\,|I_k|$.

b. $\dfrac{|f^{-2}(I_j) \cap I_i|}{|I_i|} = \displaystyle\sum_{k=1}^n \frac{|f^{-1}(I_k) \cap I_i|}{|I_i|} \frac{|f^{-1}(I_j) \cap I_k|}{|I_k|}$

(Hint: Use the fact that f is linear on each I_j with slope s_j, and $|f(b) - f(a)| = |s_j||b - a|$ if $[a, b] \subset I_j$.)

Solution: Notice first this general result. Let $J_k \subset I_k$ be any subinterval of any $I_k, 1 \leq k \leq n$. Then $f^{-1}(J_k) \cap I_i$ is a subinterval of I_i, empty interval included. Because f is linear on I_i with $|s_i| = m_i > 0$ being the absolute value of its slope on I_i, we must have

$$|f^{-1}(J_k) \cap I_i| = \frac{|J_k \cap f(I_i)|}{m_i}.$$

Either $J_k \cap f(I_i) = \varnothing$, in which case

$$|f^{-1}(J_k) \cap I_i| = |J_k \cap f(I_i)|/m_i = 0,$$

or $J_k \cap f(I_i) \neq \varnothing$, in which case $J_k \cap f(I_i) = J_k$ because f is Markovian as $I_k \subset f(I_i)$, and

$$|f^{-1}(J_k) \cap I_i| = |J_k|/m_i.$$

Now we prove (a). Two cases are considered. For the first case when either $|f^{-1}(I_k) \cap I_i| = 0$ or $|f^{-1}(I_j) \cap I_k| = 0$, equivalent to $I_k \cap f(I_i) = \varnothing$ or $I_j \cap f(I_k) = \varnothing$, we must have

$$(f^{-1}(I_j) \cap I_k) \cap I_i \subset f^{-1}(I_k) \cap I_i = \varnothing \text{ and } |f^{-1}(f^{-1}(I_j) \cap I_k) \cap I_i| = 0$$

or

$$f^{-1}(f^{-1}(I_j) \cap I_k) \cap I_i \subset f^{-1}(\varnothing) \cap I_i \text{ and } |f^{-1}(f^{-1}(I_j) \cap I_k) \cap I_i| = 0.$$

For the second case when $|f^{-1}(I_k) \cap I_i| \neq 0$ and $|f^{-1}(I_j) \cap I_k| \neq 0$, we have from the general result

$$\begin{aligned} |f^{-1}(f^{-1}(I_j) \cap I_k) \cap I_i|\,|I_k| &= |f^{-1}(J_k) \cap I_i|\,|I_k| \\ &= \frac{|J_k|}{m_i}|I_k| \\ &= |f^{-1}(I_j) \cap I_k|\frac{|I_k|}{m_i} \\ &= \frac{|I_j|}{m_k}\frac{|I_k|}{m_i} \\ &= |f^{-1}(I_j) \cap I_k|\,|f^{-1}(I_k) \cap I_i|. \end{aligned}$$

(b) Because $\{I_k : 1 \leq k \leq n\}$ is a Markovian partition of I, we have

$$\begin{aligned} \frac{|f^{-2}(I_j) \cap I_i|}{|I_i|} &= \frac{|f^{-1}(f^{-1}(I_j) \cap [\cup_{k=1}^n I_k]) \cap I_i|}{I_i} \\ &= \sum_{k=1}^n \frac{|f^{-1}(f^{-1}(I_j) \cap I_k) \cap I_i|}{I_i} \\ &= \sum_{k=1}^n \frac{|f^{-1}(I_k) \cap I_i|}{|I_i|}\frac{|f^{-1}(I_j) \cap I_k|}{|I_k|}. \end{aligned}$$

2. Let $P = [p_{ij}]$ be a probability transition matrix which is irreducible and transitive with all positive entries for P^k for some $k \geq 1$. Complete the proof of the Perron-Frobenius Theorem that $\lim_{t\to\infty} P^{kt} \to W$ implies $\lim_{t\to\infty} P^t \to W$.

Solution: Because $WP = W$, $WP^\ell = W$ for all $\ell = 0, 1, 2, \ldots, k - 1$. Thus, every sequence of the k subsequences $P^{kt+\ell}, 0 \leq \ell \leq k - 1$ converges to the same limit

$$\lim_{t\to\infty} P^{kt+\ell} = [\lim_{t\to\infty} P^{kt}]P^\ell = WP^\ell = W,$$

implying $P^n \to W$ as $k \to \infty$. This is because for every $\epsilon > 0$, there are T_ℓ so that for $t > T_\ell$, $\|P^{kt+\ell} - W\| < \epsilon$. Let $N = \max\{kT_\ell + \ell : 0 \le \ell \le k-1\}$. Then for $n > N$ there is a t and $0 \le \ell < k$ so that $n = kt + \ell$ and

$$\|P^n - W\| = \|P^{kt+\ell} - W\| < \epsilon,$$

showing what is claimed.

3. Let X be a Banach space with norm $|\cdot|$. Let $T : X \to X$ be a linear operator (with the operator norm $|T| = \sup_{|x|\le 1} T(x)$). If $|T| \le \theta < 1$, then the linear operator $I - T$ is invertible, and the inverse is

$$[I-T]^{-1} = I + T + T^2 + \cdots = \sum_{n=0}^{\infty} T^n$$

with bound

$$|[I-T]^{-1}| \le \frac{1}{1-\theta}.$$

Solution: First $A = \sum_{n=0}^{\infty} T^n$ converges uniformly because it is dominated by $\sum \theta^n$ with $\theta < 1$, and $|A| \le \frac{1}{1-\theta}$. Second, for each partial sum $A_n = \sum_{i=0}^{n} T^i$, we have

$$[I-T]A_n = [I-T][I + T + T^2 + \cdots + T^n] = I - T^{n+1}.$$

Take limit as $n \to \infty$ to get $[I-T]A = I$ as $A_n \to A$ and $T^n \to 0$ since $\|T^n\| \le \theta^n$. Similarly, we can have $A[I-T] = I$, showing $[I-T]^{-1} = A$ as required.

4. Let X, Y be Banach spaces, and denote by $L(X, X)$ the Banach space of all bounded linear operators from X to X with the operator norm defined as in Problem 3 above. Let $T : Y \to L(X, X)$. Assume there is a constant $\theta < 1$ so that $\sup_{y\in Y} |T(y)| \le \theta < 1$, and T is continuously differentiable in y, i.e., for $\Delta = T(y+h) - T(y) - DT(y)h$, $|\Delta| = o(|h|)$, and $DT(y) \in L(Y, L(X, X))$ is continuous. Then $[I - T(y)]^{-1}$ is also continuously differentiable in y. Describe how the result can be generalized to the case that $T(y)$ is C^k or analytic in y. (Hint: Show first the power rule $DT^n(y)h = [DT(y)h]T^{n-1}(y) + T(y)[DT(y)h]T^{n-2}(y) + \cdots + T^{n-1}(y)[DT(y)h]$ for $y, h \in Y$. Also, you can cite without proof any reasonable result on uniformly convergent series.)

Solution: For $y, h \in Y$, we have

$$\begin{aligned}
\|\Delta_n\| &:= \|T(y+h)^n - T(y)^n - DT^n(y)h\| \\
&= \|(T(y+h) - T(y) - DT(y)h)T(y+h)^{n-1} \\
&\quad + T(y)T(y+h)^{n-1} + (DT(y)h)[T(y+h)^{n-1} - T(y)^{n-1}] - T(y)^n \\
&\quad - (T(y)[DT(y)h]T^{n-2}(y) + T(y)^2[DT(y)h]T^{n-3}(y) + \cdots + T^{n-1}(y)[DT(y)h]\| \\
&\le \|\Delta\|\|T(y+h)^{n-1}\| + \|DT(y)h\|\|T(y+h)^{n-1} - T(y)^{n-1}\| \\
&\quad + \|T(y)\|\|T(y+h)^{n-1} - T(y)^{n-1} - ([DT(y)h]T^{n-2}(y) + \cdots + T^{n-2}(y)[DT(y)h])\|
\end{aligned}$$

Notice that the first two terms are $o(\|h\|)$. For the third term, we can apply the same technique recursively to show $\|\Delta_n\|$ is $o(\|h\|)$. So $T(y)^n$ is differentiable with the given derivative. Moreover, the derivative is bounded by

$$\|DT^n(y)\| \le n\|DT\|\|T\|^{n-1} \le n\theta^{n-1}\|DT\|.$$

Hence, the infinite series of the derivatives is dominated by $\sum n\theta^{n-1}\|DT\|$ and therefore converges to the derivative of the sum, i.e., $[I - T(y)]^{-1}$ is differentiable in y. This argument can be used to show $[I - T(y)]^{-1}$ is C^k if $T(y)$ is C^k. If $T(y)$ is analytic, so is $[I - T(y)]^{-1}$ because it is differentiable in the complex variable y.

Alternate: Prove first the product rule: Let $h(x) = f(x)g(x)$ where $f, g \in C^1(U, L(Y, Y))$. U open in X, and X, Y are Banach spaces. Then for $x \in U$ and $h \in X$, $Dh(x)h = [Df(x)h]g(x) + f(x)[Dg(x)h]$.

(f is not linear in x because U may just be a small open set, but $f(x), x \in U$ is linear in Y, i.e., $f(x)(ay + bz) = af(x)(y) + bf(x)(z)$ for any $y, z \in Y$ and $a, b \in \mathbb{R}$. Also, $h(x)(y) = f(x)(g(x)(y))$ for any $y \in Y$ and $x \in X$.)

Proof of the Product Rule: Let $\Delta = h(x+h) - h(x) - Dh(x)h$, then we have by inserting four auxiliary terms

$$\Delta = h(x+h) - f(x)g(x+h) + f(x)g(x+h) - [Df(x)h]g(x+h) + [Df(x)h]g(x+h) - h(x) - Dh(x)h.$$

Regroup the terms to have

$$\begin{aligned}
\|\Delta\| &\le \|(f(x+h) - f(x) - [Df(x)h])g(x+h)\| + \|[Df(x)h](g(x+h) - g(x)\| \\
&\quad + \|f(x)(g(x+h) - g(x) - [Dg(x)h])\|.
\end{aligned}$$

For which, the first term is of order $o(h)$ because f is differentiable, the second order is of $o(h)$ because $g(x+h)-g(x) \to 0$ and $\|[Df(x)h]\| \leq \|[Df]\|_0\|h\|$, and the third term is of $o(h)$ because g is differentiable. Thus, by definition, h is differentiable at $x \in U$ with the said derivative, completing the proof for the product rule.

As an application, the power-rule follows as below:

$$\begin{aligned} D(T^n(y))h &= [DT(y)h]T(y)^{n-1} + T(y)[D(T(y)^{n-1})h] \\ &= [DT(y)h]T(y)^{n-1} + T(y)([DT(y)h]T(y)^{n-2} + T(y)[D(T(y)^{n-2})h]) \\ &= \dots \\ &= [DT(y)h]T^{n-1}(y) + T(y)[DT(y)h]T^{n-2}(y) + \cdots + T^{n-1}(y)[DT(y)h] \end{aligned}$$

5. Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ be a C^k function with $k \geq 1$. Assume $x_0 \in \mathbb{R}^n$ is a hyperbolic fixed point of $f(\cdot, y_0)$ for some y_0. Prove that x_0 persists in the sense that there is a small neighborhood V of y_0 and U of x_0 and a C^k function $\phi : V \to U$ so that $f(x, y) = x$ in $U \times V$ iff $x = \phi(y)$.

Solution: Use Implicit Function Theorem for $F(x, y) = x - f(x, y)$, and use the fact that $D_xF(x_0, y_0) = I - D_xf(x_0, y_0)$ is nonsingular, thus invertible, because $\lambda = 1$ is not an eigenvalue of $D_xf(x_0, y_0)$.