Prof. McCann Fields Academy / MAT 1502 Assignment 1 Due: Monday Jan 24

This assignment is intended primarily (a) to help students understand the course material and (b) to help the instructor understand the level and background of the students. Do not be too concerned if you find some problems difficult; just give them your best try.

1. **Riemannian geodesics and distance.** (a) Let $g_{ij}(x)$ be a smooth map from $\mathbf{R}^n$ into symmetric positive definite $n \times n$ matrices, and fix $x_0, x_1 \in \mathbf{R}^n$. If $x : [0,1] \to \mathbf{R}^n$ minimizes the energy functional
$$E[x] := \frac{1}{2} \int_0^1 g_{ij}(x(t)) \dot{x}^i(t) \dot{x}^j(t) dt$$
among all smooth curves starting at $x(0) = x_0$ and ending at $x(1) = x_1$, use calculus to show that $x(t)$ satisfies the ordinary differential equation
$$\ddot{x}^i(t) + \Gamma^i_{jk}(x(t)) \dot{x}^j(t) \dot{x}^k(t) = 0,$$
where
$$\Gamma^i_{jk} = \frac{1}{2} g^{im} \left(\frac{\partial g_{mk}}{\partial x^j} + \frac{\partial g_{jm}}{\partial x^k} - \frac{\partial g_{jk}}{\partial x^m} \right)$$
are the so-called Christoffel symbols. (Repeated indices are summed from 1 to n, and g^{im} denotes the matrix inverse of g_{mj}.)

b) In that case, show (using Jensen's inequality, for example) that $x(t)$ also minimizes the length functional
$$L[x] := \int_0^1 (g_{ij}(x(t)) \dot{x}^i(t) \dot{x}^j(t))^{1/2} dt$$
in the same class of curves. Its value at the miminum defines $d(x_0, x_1) := L[x]$.

c) Using (a), show by induction that if $x : [0,1] \longrightarrow \mathbf{R}^n$ minimizes $E[x]$ in the larger class of continuous, piecewise smooth curves, then $x \in C^k$ for all $k \in \mathbf{N}$.

d) Show $d(x_0, x_1) \le d(x_0, y) + d(y, x_1)$ for all $y \in \mathbf{R}^n$ and $d(x(s), x(t)) = |s - t| d(x_0, x_1)$ for all $s, t \in [0,1]$.

2. **Functional analysis.**
Give a statement of the (a) Riesz-Markov and (b) Banach-Alaoglu theorems.

(c) Given a pair of Borel probability measures $\mu^\pm$ on a compact separable metric space (X, d), use them to show the set of non-negative joint measures $\Gamma(\mu^+, \mu^-)$ on $X \times X$ having μ^+ and μ^- for marginals is weak-$*$ compact.

(d) Let $c \in C(X)$ be a continuous function on the same space X. Show the cost functional
$$cost(\gamma) := \int_{X^2} c(x,y) d\gamma(x,y)$$
attains its minimum on $\Gamma(\mu^+, \mu^-)$.

3. **Doubly stochastic matrices.** (a) Give a statement of the Krein-Milman theorem.

(b) Use it to show that at least one of the minimizers in problem #2 above is an extreme point, meaning it fails be the midpoint of any segment in $\Gamma = \Gamma(\mu^+, \mu^-)$.

(c) Let $X = \{1, 2, \ldots, n\}$ and $\mu^\pm = \frac{1}{n} \sum_{i=1}^n \delta_i$. Show the set $\Gamma(\mu^+, \mu^-)$ corresponds to the set of non-negative $n \times n$ matrices whose columns and rows each sum to 1.

(d) In case (c) show that Γ has precisely $n!$ extreme points.

1. (a) We can use the Euler-Lagrange equations:

$$\frac{\partial L}{\partial x^k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^k} = 0$$

$$\frac{\partial(\dot{x}^i \dot{x}^j)}{\partial \dot{x}^k} = \begin{cases} 0 & \text{if } i \neq k, j \neq k \\ \dot{x}^i & \text{if } j = k, i \neq k \\ \dot{x}^j & \text{if } i = k, j \neq k \\ 2\dot{x}^k & \text{if } i = j = k \end{cases}$$

therefore $\frac{\partial L}{\partial \dot{x}^k} = g_{ik}\dot{x}^i$

Now $\frac{d}{dt}\frac{\partial L}{\partial \dot{x}^k} = \frac{\partial g_{ik}}{\partial x^j}\dot{x}^i\dot{x}^j + g_{ik}\ddot{x}^i$

Rewrite the first term as $\frac{1}{2}\frac{\partial g_{ik}}{\partial x^j}\dot{x}^i\dot{x}^j + \frac{1}{2}\frac{\partial g_{ik}}{\partial x^j}\dot{x}^i\dot{x}^j$

$$= \frac{1}{2}\frac{\partial g_{ik}}{\partial x^j}\dot{x}^i\dot{x}^j + \frac{1}{2}\frac{\partial g_{jk}}{\partial x^i}\dot{x}^i\dot{x}^j$$

with interchanged indices in the last term.

$$\frac{\partial L}{\partial x^k} = \frac{1}{2}\frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j$$

So if we put it together we get

(*) $$g_{ik}\ddot{x}^i + \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right)\dot{x}^i\dot{x}^j = 0$$

$k = 1, \ldots, n$

Denote $\Gamma_{kij} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right)$

Then we can write

$$(g_{ij})\begin{pmatrix} \ddot{x}^1 \\ \vdots \\ \ddot{x}^n \end{pmatrix} = -\begin{pmatrix} \Gamma_{1ij}\dot{x}^i\dot{x}^j \\ \vdots \\ \Gamma_{nij}\dot{x}^i\dot{x}^j \end{pmatrix}$$

And if we define $(g^{hk}) = (g_{ij})^{-1}$ and $\Gamma^k_{ij} = g^{hk}\Gamma_{hij}$

then we can rearrange:

$$\begin{pmatrix} \ddot{x}^1 \\ \vdots \\ \ddot{x}^n \end{pmatrix} = -(g^{hk})\begin{pmatrix} \Gamma_{1ij}\dot{x}^i\dot{x}^j \\ \vdots \\ \Gamma_{nij}\dot{x}^i\dot{x}^j \end{pmatrix} = -\begin{pmatrix} \Gamma^1_{ij}\dot{x}^i\dot{x}^j \\ \vdots \\ \Gamma^n_{ij}\dot{x}^i\dot{x}^j \end{pmatrix}$$

So $\ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0$, $k = 1, \ldots, n$.

1. (b) Let $f = (g_{ij}(x(t))\dot{x}^i\dot{x}^j)^{1/2}$. $f \in L^2([0,1])$ since it is cts.
Then by Cauchy-Schwarz,

$$L(x)^2 = \left|\int_0^1 f \cdot 1 \, dt\right|^2 \leq \left(\int |f|^2\right)\left(\int 1^2\right)$$
$$= 2E(x)$$

We have equality iff f and 1 are linearly dependent i.e. f is constant a.e.

Let x minimize E and y be another candidate. Then

$$L(x)^2 \leq 2E(x) \leq 2E(y^*) = L(y^*)^2 = L(y)^2$$

since x minimizes E where y^* is the arc length parameterization of y

It remains to be shown that length is independent of parametrization. Let φ be a change of parameter. Then

$$L(y \circ \varphi) = \int_0^1 \left(g_{ij}(y(\varphi(t)))\dot{y}^i(\varphi(t))\,\dot{y}^j(\varphi(t))\left(\frac{d\varphi}{dt}\right)^2\right)^{1/2} dt$$
$$= L(y) \text{ by change of variables}$$

Therefore $L(x) \leq L(y)$.

(c) Suppose x minimizes $E(x)$ in the class of cts piecewise smooth curves and is C^{k-1}. Then x is C^k.

We just need to show that $\frac{d^{k-1}x^i}{dt^{k-1}}$ has cts derivative at a point where it is a priori only continuous.

Using the equation from (a), we have a formula for $\ddot{x}^i$ which is smooth in t. By differentiating this equation wrt t or integrating we obtain that $\frac{d^k x^i}{dt^k}$ is a smooth function of t and lower order partials.

Let $x:[a,b]\to\mathbb{R}^n$ and $y:[c,d]\to\mathbb{R}^n$ which are smooth and are pieced together as part of a minimizing curve. Then they must individually satisfy the equation for (a). Let the full minimizing curve $z:[0,1]\to\mathbb{R}^n$ and let $x:[a,b]\subset[0,1]\to\mathbb{R}^n$ and $y:[b,c]\to\mathbb{R}^n$. (A)

Let $\frac{d^{k-1}z^i}{dt^{k-1}}(t)=:f^i(t)$. We want to show that f^i has cts derivative, on (a,c). Since the equation from (a) applies on each piece, $f^{i\prime}(t)$ is defined everywhere

and cts except at $t=b$. But the limit $\lim_{t\to b} f^{i\prime}(t)$ exists since the equation gives it as a smooth function of t and lower order partials and $z\in C^{k-1}$. Therefore since f^i is cts, it is well-known that f^i is differentiable at b with derivative $\lim_{t\to b} f^{i\prime}(t)$. Therefore f^i has cts derivative on (a,c). This may be applied at every such point to obtain that $z\in C^k$.

Therefore $z\in C^\infty$. (For the case $k=1$, $f^{i\prime}$ depends on $\int h(t, x(t), \dot{x}(t))$ but this is okay because h is Lebesgue integrable since it is a smooth function of cts or bounded functions so $\int h$ is cts and does not depend on $\dot{x}(b)$ since singletons are null. Then the limit still exists.)

(d) We have $d(x_0,y)=L(a)$, $d(y,x_1)=L(b)$ and $d(x_0,x_1)=L(c)$ for some smooth curves a,b,c. But a concatenated with b is in the same class as c since (c) shows that a concatenated with b is smooth. So $L(c)\leq L(a)+L(b)$ since c minimizes L and L is invariant under reparametrization and the length of a concatenation is the sum of the lengths. Therefore d satisfies the triangle inequality.

We claim that the integrand in L is constant for the minimizer.

$$\frac{d}{dt}\left(g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t)\right)$$

$$= g_{ij}\ddot{x}^i\dot{x}^j + g_{ij}\dot{x}^i\ddot{x}^j + \frac{\partial g_{ij}}{\partial x^k}\dot{x}^i\dot{x}^j\dot{x}^k$$

by (*) in (a)

$$= -\left(\frac{\partial g_{ik}}{\partial x^l} + \frac{\partial g_{li}}{\partial x^k} - \frac{\partial g_{lk}}{\partial x^j}\right)\dot{x}^l\dot{x}^k\dot{x}^j + \frac{\partial g_{lj}}{\partial x^k}\dot{x}^k\dot{x}^l\dot{x}^j$$

$= 0$ by rearranging indices. ///.

Then

$$d(x(s),x(t)) = \int_{\min(s,t)}^{\max(s,t)} \underbrace{\left(g_{ij}(x(\tau))\dot{x}^i\dot{x}^j\right)^{1/2}}_{=c} d\tau$$

since x must minimize L from $x(s)$ to $x(t)$ otherwise we could replace it and it would be smooth by part (c) and contradict optimality of x

$$= |s-t|c$$

and $d(x_0,x_1) = \int_0^1 c = c$

so $d(x(s),x(t)) = |s-t|\,d(x_0,x_1)$.

2. (a) Let X be locally compact Hausdorff and for $\mu \in M(X)$ and $f \in C_0(X)$ let $I_\mu(f) = \int f d\mu$. Then $\mu \to I_\mu$ is an isometric isomorphism from $M(X)$ to $C_0(X)^*$.

(b) If X is a normed space, the closed unit ball in X^* is weak $*$ compact.

(c) Since X^2 is compact Hausdorff, Riesz-Markov implies that we may identify $C(X^2)^*$ and $M(X^2)$. Since $\Gamma(\mu^+, \mu^-)$ consists of Borel probability measures and X^2 is second countable, $\Gamma(\mu^+, \mu^-) \subset M(X^2)$. Furthermore, probability measures are a closed subset of the unit ball in $M(X^2)$, since X^2 is compact.

Therefore, by Banach-Alaoglu it suffices to show that $\Gamma(\mu^+, \mu^-)$ is closed. Let $\pi_1 : P(X^2) \to P(X)$ take Borel probability measures on X^2 into those on X by $\pi_1(\mu)(A) = \mu(A \times X)$. Equip these spaces with the subspace topology from the weak $*$ topology on $M(\mathbb{X})$, the corresponding space of Radon measures. Define π_2 similarly.

Then $\Gamma(\mu^+, \mu^-) = \pi_1^{-1}(\{\mu^+\}) \cap \pi_2^{-1}(\{\mu^-\})$. It therefore suffices to show π_i are cts since $P(X)$ is metrizable and thus Hausdorff since X is separable.

Let $\Sigma \subset P(X)$ be closed i.e. if $\mu_i(f) \to \mu(f)$ $\mu_i \in \Sigma$ $\forall f \in C(X)$ then $\mu \in \Sigma$ since $P(X)$ is first countable since X is separable. Let $\mu_j \in \pi_1^{-1}(\Sigma)$. If $\mu_j(f) \to \mu(f)$ $\forall f \in C(X^2)$ we know that $(\pi_1(\mu_j))(f^*) \to (\pi_1(\mu))(f^*)$ where $f^*(x) = f(x, y_0)$ for some arbitrary choice of y_0 since $(\pi_1(\mu^*))(f^*) = \mu^*(f^*)$ where f^* in the second instance is viewed as a function of two variables and $f^* \in C(X)$, where $\mu^* = \mu_j, \mu$.

But $\pi_1(\mu_j) \in \Sigma$ and the mapping $f \mapsto f^*$ is surjective onto $C(X)$ so $\pi_1(\mu) \in \Sigma$. Therefore $\mu \in \pi_1^{-1}(\Sigma)$ and π_1 is cts. The same argument applies to π_2. So $\Gamma(\mu^+, \mu^-)$ is weak $*$ compact.

(d) cost : $M(X^2) \to \mathbb{C}$ is a linear functional. It is bounded and thus cts because the weak * topology guarantees all such evaluation maps are continuous: $\text{cost}(\gamma) = \langle c, \gamma \rangle$ where $c \in C(X^2)$ and $\gamma \in M(X^2)$ is a dual pairing.

Then cost is a cts function on a compact set so its image is compact. If we assume c is a real-valued function then on $\Gamma(\mu^+, \mu^-)$ it is real-valued so it attains its minimum.

3. (a) A compact convex subset of a Hausdorff locally convex topological vector space is equal to the closed convex hull of its extreme points.

(b) $\Gamma(\mu^+, \mu^-)$ is compact convex in $M(X^2)$, and in $P(X^2)$. Let $A = \{\mu \in \Gamma(\mu^+, \mu^-) \mid \text{cost}(\mu) = \min_{\Gamma(\mu^+,\mu^-)} \text{cost}\}$. Since cost

is cts, A is closed and therefore compact. It is also easily seen that it is convex: let $\mu, \nu \in A$, $\mu \neq \nu$, $a \in (0,1)$.
$\text{cost}(a\mu + (1-a)\nu) = a\,\text{cost}(\mu) + (1-a)\,\text{cost}(\nu) = \min_{P(\mu^+,\mu^-)} \text{cost}$.
since cost is linear so A is convex.

So A is compact convex in $P(X^2)$ so by Krein-Milman it contains an extreme point. If we can show that an extreme point of A is an extreme point of $\Gamma(\mu^+, \mu^-)$, we are done.

This is true because A is a face: if $a \neq b \in \Gamma(\mu^+, \mu^-)$ and $aa + (1-a)b \in A$, $a \in (0,1)$ then $a, b \in A$. This is also easy to show: $\text{cost}(a) = \text{cost}(b) = \min_{\Gamma(\mu^+,\mu^-)} \text{cost}$ by the linearity of cost.

Therefore, if a point is not an extreme point of $\Gamma(\mu^+, \mu^-)$ but is in A it is not an extreme point of A.
Therefore, at least one minimizer is an extreme point.

3. (c) Establish a bijection $f: \Gamma(\mu^+, \mu^-) \to M_n = \{\text{the given set}\}$ by $f(\mu)_{ij} = n\mu(i \times j)$. Since $\sum \mu(i \times j) = \mu(i \times X) = \frac{1}{n}$, the columns and rows sum to 1.
The inverse is $f^{-1}(M)(i \times j) = \frac{M_{ij}}{n}$. This defines a unique measure since the space is finite.

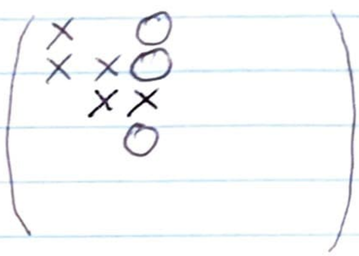
(d) We claim that the extreme points of Γ are precisely
θ those which correspond to $n \times n$ matrices with four entries with 1 and the rest zeroes in M_n. It is easy to compute that there are $n!$ of these since there is n positions to place the one in the first column, $n-1$ to place the second since one row is disallowed and so on.

It remains to show the first statement. It is clear that these elements are extreme. For it to be a midpoint, the endpoints would have to be "more extreme" in some position/on some set. This is not possible since no element of M_n has an entry outside the set $[0,1]$.

To prove the converse, suppose an element μ of Γ not in this
θ set is extreme. We construct a line segment which μ lies in the middle of. We work with the matrix representation which preserves the algebraic structure. M_μ must contain
i,j an entry with values in $(0,1)$ since it is not in the stated set, θ. Its column must then contain another such entry since columns sum to 1, (i', j), and so on to (i', j'). Now if (i, j') satisfies this, we are done. Otherwise, we proceed to (i'', j'). We claim there always exists a set of pairs of indices A and B which are disjoint with at most one per column and row and one of each when one is present. These correspond to entries we can perturb.
For example, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \boxed{\frac{1}{2}} \\ 0 & \boxed{\frac{1}{2}} & \frac{1}{2} \end{pmatrix}$ Indeed, ignoring columns and rows of 1 entries, we have a $m \times m$ matrix where $n \geq m \geq 2$.

In this reduced matrix, every column and row contains at least two nonzero entries since the columns and rows sum up to 1 and no entries are equal to 1. We claim there exists a finite sequence of entries (i_p, j_p) which are nonzero and distinct such that consecutive terms only change one component in an alternating fashion and each row/column contains either zero or two of these terms. We will try to place nonzero entries pathologically so they do not satisfy this. We start with $(1,1), (2,1), (2,2)$ (wlog by renumbering).

<u>Claim</u> At each subsequent stage, to avoid creating such a finite sequence, the entry must be placed in a new column or row. After this, all columns either have zero or two such entries and same for all rows except the first and perhaps last, which can have one.
Suppose it is time to change the row (as above). The full sequence would be a path if we place in the first row. Placing in a row previously traversed with two entries would create a path starting from the second entry in that row. So to avoid this, we must go to a new row. The same argument holds when it's time to change the column. As a result, the second sentence holds.



But the matrix only has m rows and columns so such a path must exist. Let $e_{ij} = \min(|M_{ij} - 1|, |M_{ij} - 0|)$. Let ε be the min of e_{ij} over the elements in the path, $\varepsilon > 0$. Now we add ε to each odd entry in the sequence and subtract ε from each even entry, to obtain $\nu \in \Gamma$ and reverse signs to obtain $\sigma \in \Gamma$. But $\mu = \frac{1}{2}\nu + \frac{1}{2}\sigma$ so μ is not extreme. So Γ has precisely $n!$ extreme points.