MAT 1502: Assignment #3

Due: noon on Tuesday, Apr 5

Also: draft project due Tuesday March 22; final version due Thursday 7 April (midnight)

1. A positive-definite symmetric matrix Σ is called the *covariance* of the centered ellipsoid
$$E := \{x \in \mathbf{R}^n \mid \langle x\Sigma^{-1}x\rangle \leq 1\}.$$
Given uniform probability measures on m different centered ellipsoids in $\mathbf{R}^n$, show the (equally weighted) 2-Wasserstein barycenter μ of $\mu_1, \ldots, \mu_m$ is also uniform on some centered ellipsoid and identify its covariance Σ in terms of $\Sigma_1, \ldots, \Sigma_m$.

2. Fix $N > 0$, $K \in \mathbf{R}$, $\phi \in C^2(\mathbf{R}^n)$ and let $d(x,y) = |x-y|$ and vol denote the Euclidean distance and volume on $M := \mathbf{R}^n$. Setting $dm(x) = e^{-\phi(x)} d\,\mathrm{vol}(x)$, we've seen that $(M,d,m) \in CD^e(K,N)$ if $N = n$ and $\phi = \mathit{const}$ and $K \leq 0$, or if $N > n$ and
$$D^2\phi - \frac{1}{N-n} D\phi \otimes D\phi - KI \geq 0 \qquad \forall x \in \mathbf{R}^n \qquad (1)$$
is non-negative definite. Show the converse by constructing d_2-geodesics along which $(Kd_2(\mu_0,\mu_1)^2, N)$ convexity of the entropy fails if (a) $N > n$ but (1) is violated for some $x \in \mathbf{R}^n$ or if (b) $N = n$ but ($\phi \neq \mathit{const}$ or $K > 0$) or if (c) $0 < N < n$.

3. (a) Given functions $V_1 : M \longrightarrow (-\infty, \infty]$ and $V_2 : M \longrightarrow (-\infty, \infty]$ on a metric space (M,d) with V_i is weakly (K_i, N_i)-convex for $i \in \{1,2\}$ and V_2 strongly (K_2, N_2)-convex, show $V := V_1 + V_2$ is weakly $(K_1 + K_2, N_1 + N_2)$-convex.
(b) Given $(M,d,m) \in CD^e(K_1, N_1)$ with $m(M) < \infty$ and $V : M \longrightarrow [0,\infty]$ Borel and strongly (K_2, N_2)-convex, show $(M, d, e^{-V}m) \in CD^e(K_1 + K_2, N_1 + N_2)$.

4.(Γ-convergence). For $k \in \mathbf{N}_\infty := \mathbf{N} \cup \{\infty\}$, let $f_k : M \longrightarrow (-\infty, \infty]$ be a sequence of lower semicontinuous functions on a compact metric space (M,d). We say f_k has Γ-limit f_∞ if the following two criteria are satisfied: (i)
$$f_\infty(x_\infty) \leq \liminf_{k\to\infty} f_k(x_k) \qquad \text{whenever } 0 = \lim_{k\to\infty} d(x_k, x_\infty);$$
and (ii) every $y_\infty \in M$ is the limit of a sequence $y_k \in M$ along which
$$f_\infty(y_\infty) = \lim_{k\to\infty} f_k(y_k).$$

(a) Show that if f_k has Γ-limit f_∞ then any sequence $x_k \in \arg\min f_k$ has a subsequence whose limit x_∞ satisfies $x_\infty \in \arg\min f_\infty$.
(b) Give an example of Γ-convergence in which there are elements of $\arg\min f_\infty$ not approximated by any such sequence.

MAT1502 PS3

Jeffrey Liang

March 2022

1. We follow Agueh and Carlier (2011) who prove an analoguous result for Gaussians. We claim that the barycenter μ is a centered ellipsoid with covariance Σ which is the unique positive definite solution to the matrix equation:

$$\sum_{i=1}^{m} \lambda_i (\Sigma^{1/2} \Sigma_i \Sigma^{1/2})^{1/2} = \Sigma \tag{1}$$

First let us remark that linear transformations by a positive definite matrix take uniform measures on centered ellipsoids to uniform measures on centered ellipsoids. If we have a centered ellipsoid $A := \{x | x\Sigma^{-1}x \leq 1\}$ and a positive definite matrix B, then $BA = \{x | xB^{-1}\Sigma^{-1}B^{-1}x \leq 1\}$, a centered ellipsoid with covariance $B\Sigma B$. Moreover, these maps will have constant Jacobian determinant $detB$ so the resulting measure is still uniform. Also since the linear map is the gradient of a convex function $x \to \frac{1}{2}xBx$, it is the optimal map for the squared distance.

Now we show the existence of a positive definite solution to (1). Let α_i, β_i denote respectively the smallest and largest eigenvalue of Σ_i and let α, β be such that

$$\beta \geq (\sum_{i=1}^{m} \lambda_i \sqrt{\beta_i})^2 \geq (\sum_{i=1}^{m} \lambda_i \sqrt{\alpha_i})^2 \geq \alpha. \tag{2}$$

If we define $K_{\alpha,\beta}$ to be the set of symmetric nxn matrices Σ such that $\beta I \geq \Sigma \geq \alpha I$, we note that this set is convex and compact. If we define $F(\Sigma) := \sum_{i=1}^{m} \lambda_i (\Sigma^{1/2}\Sigma_i\Sigma^{1/2})^{1/2}$ on $K_{\alpha,\beta}$, we note that this is a continuous self-map: by square rooting and multiplying by $\sqrt{\alpha}, \sqrt{\beta}$,

$$\beta I \geq \sum_{i=1}^{m} \lambda_i \sqrt{\beta\beta_i} I \geq F(\Sigma) \geq \sum_{i=1}^{m} \lambda_i \sqrt{\alpha\alpha_i} I \geq \alpha I \tag{3}$$

for all $\Sigma \in K_{\alpha,\beta}$. Brouwer's fixed point theorem then shows the existence of a solution Σ.

For any positive definite solution Σ of (1), if we set μ to be the uniform measure on the centered ellipsoid with covariance Σ, we have that the optimal transport from μ_i to μ is given by the positive definite linear map $T_i = \Sigma_i^{1/2}(\Sigma_i^{1/2}\Sigma\Sigma_i^{1/2})^{-1/2}\Sigma_i^{1/2}$. This can be easily checked, as from the above discussion, the covariance of the centered ellipsoid which is the support of the image measure is $T_i\Sigma_i T_i = \Sigma$. Agueh and Carlier show that it is sufficient that $\sum_{i=1}^{m} \lambda_i T_i = I$ for μ to be the barycenter.

Set $D_i := \Sigma_i^{1/2}$ and $K := \Sigma^{1/2}$. With the identity

$$(KD_i^2K)^{1/2} = KD_i(D_iK^2D_i)^{-1/2}D_iK, \tag{4}$$

we can rewrite $F(\Sigma) = \Sigma$ as

$$\sum_{i=1}^{m} \lambda_i KD_i(D_iK^2D_i)^{-1/2}D_iK = K^2 \tag{5}$$

but K is invertible so we can simplify and obtain

$$\sum_{i=1}^{m} \lambda_i D_i(D_iK^2D_i)^{-1/2}D_i = \sum_{i=1}^{m} \lambda_i T_i = I, \tag{6}$$

so μ is indeed the barycenter.

Finally, we have shown that for any positive definite solution Σ of (1), it is the covariance of the centered ellipsoid which is the support of the barycenter. However, the barycenter is unique so there must be a unique positive definite solution of (1).

2. (a) We showed in class that

$$\frac{1}{N}(e'(0))^2 \leq \frac{1}{N}\int_M (D\phi Du - \Delta u)^2 d\mu_0 \tag{7}$$

$$\leq \frac{1}{N}\int_M (1+\epsilon^{-1})(D\phi Du)^2 + (1+\epsilon)nTr((D^2u)^2)d\mu_0 \tag{8}$$

$$\leq \int_M \frac{1}{N-n}Du(D\phi \otimes D\phi)Du + Tr((D^2u)^2)d\mu_0 \tag{9}$$

$$\leq \int_M (DuD^2\phi Du + Tr((D^2u)^2) - KDu \cdot Du)d\mu_0 \tag{10}$$

$$= e''(0) - Kd_2(\mu_0, \mu_1)^2 \tag{11}$$

where $e(t)$ is the entropy at time t and $Du(x) = DU(x) - x$ where $DU(x)$ is the optimal map. (7) follows by Jensen, (8) follows whenever $\epsilon > 0$, (9) follows from a choice of $\epsilon = \frac{N-n}{n}$ and (10) follows from positive definiteness. Our strategy will be to construct $\mu_0(r), \mu_1(r)$ as functions of a real parameter $r > 0$ so that as $r \to 0$, the approximation error in the above chain of inequalities goes to 0 so that the failure of the last inequality (10) will dominate and convexity of the entropy will fail.

Let x^*, v be such that $v(D^2\phi(x^*) - \frac{1}{N-n}D\phi(x^*) \otimes D\phi(x^*) - KI)v = k < 0$. By negating v if necessary, we may assume that $D\phi(x^*)v = l \leq 0$. We can construct a u which is convex and satisfies $Du(x^*) = v$ and $\Delta u(x^*) = -\frac{n}{N-n}l$. Explicitly, let $u(x) = \sum_i a_i|x_i - x_i^*|^2 + v \cdot x$, where $a_i \geq 0$ and $\sum_i 2a_i = -\frac{n}{N-n}l$, which is possible since $l \leq 0$.

Let $\mu_0(r)$ be the uniform measure on a ball of radius r around the point x^* and $\mu_1(r)$ be the pushforward of $\mu_0(r)$ via $DU(x) := Du(x) + x$, where $u(x)$ was constructed above. Note that $U(x)$ is convex so that $DU(x)$ is d_2 optimal.

The quadratic form is continuous and takes the value of $k < 0$ at $x = x^*$. We may therefore pick δ so that on the corresponding ball around x^*, the quadratic form is bounded above by a negative number, k^*.

Let $\epsilon = -k^*$. We wish to show that for r small enough, the failure of the positive definiteness dominates the approximation error, i.e. the approximation error is less than ϵ.

The approximation error in the first inequality (7) is known as the Jensen gap and can be bounded by the variance of the integrand for the quadratic. By picking δ' sufficiently small, we may make $||(D\phi Du - \Delta u)(x) - (1 + \frac{n}{N-n})l|| < \sqrt{N\epsilon/3}$ when $||x - x^*|| < \delta'$. We can then bound the variance of $(D\phi Du - \Delta u)$ on this ball by $(\sqrt{N\epsilon/3})^2 = N\epsilon/3$. Therefore, we can make the Jensen gap arbitrarily small by picking δ' sufficiently small.

We need to deal with the approximation error of (8). A simple computation shows that this error is equal to $\frac{1}{N}\int_M(\sqrt{\frac{1}{\epsilon}}D\phi Du + \sqrt{\epsilon}\Delta u)^2 d\mu_0 = \frac{1}{N}\int_M(\sqrt{\frac{n}{N-n}}D\phi Du + \sqrt{\frac{N-n}{n}}\Delta u)^2 d\mu_0$ as a result of our choice of $\epsilon = \frac{N-n}{n}$. The integrand is a continuous function and we chose u precisely so that it is 0 at x^*. Therefore we may bound the integrand by $N\epsilon/3$ by choosing δ'' sufficiently small.

Putting things together, if we choose $r = min(\delta, \delta', \delta'')$, we have:

$$
\begin{aligned}
e''(0) - Kd_2(\mu_0,\mu_1)^2 - \frac{1}{N}(e'(0))^2 &= \int_M (DuD^2\phi Du + Tr((D^2u)^2) - KDu\cdot Du)d\mu_0 - \frac{1}{N}(e'(0))^2 && (12)\\
&= \int_M (DuD^2\phi Du + Tr((D^2u)^2) - KDu\cdot Du)d\mu_0 - (\int_M \frac{1}{N-n}Du(D\phi\otimes D\phi)Du + \\
&\quad Tr((D^2u)^2)d\mu_0) + (\int_M \frac{1}{N-n}Du(D\phi\otimes D\phi)Du + Tr((D^2u)^2)d\mu_0) - \frac{1}{N}(e'(0))^2 && (13)\\
&\le -\epsilon + (\int_M \frac{1}{N-n}Du(D\phi\otimes D\phi)Du + Tr((D^2u)^2)d\mu_0) - (\frac{1}{N}\int_M (D\phi Du - \Delta u)^2 d\mu_0) \\
&\quad + (\frac{1}{N}\int_M (D\phi Du - \Delta u)^2 d\mu_0) - \frac{1}{N}(e'(0))^2 && (14)\\
&\le -\epsilon + \epsilon/3 + (\frac{1}{N}\int_M (D\phi Du - \Delta u)^2 d\mu_0) - \frac{1}{N}(e'(0))^2 && (15)\\
&\le -\epsilon + 2\epsilon/3 && (16)\\
&< 0 && (17)
\end{aligned}
$$

(b) Let μ_0 be the uniform measure (with respect to Lebesgue measure) on the ball of radius r around the point x^* and μ_1 the uniform measure on the ball of radius R around the point $x^* + t$. Then the map that pushes forward μ_0 to μ_s on the geodesic connecting μ_0 and μ_1 is $G_s(x) = (1-s(1-R/r))x + s(x^*(1-R/r)+t)$ with Jacobian determinant $(1-s(1-R/r))^n$. Now let us compute the entropy $S(\mu_s)$:

$$
\begin{aligned}
S(\mu_s) &= \int_M \log\frac{d\mu_s(x)}{dm}d\mu_s && (18)\\
&= \int_M \log\frac{d\mu_s(G_s(x))}{dm}d\mu_0 && (19)\\
&= \int_M \log(\frac{d\mu_s(G_s(x))}{dH^n}/\frac{dm(G_s(x))}{dH^n})d\mu_0 && (20)\\
&= \int_M \log\frac{e^{\phi(G_s(x))}}{\omega(r,n)(1-s(1-R/r))^n}d\mu_0 && (21)\\
&= \int_M \phi(G_s(x))d\mu_0 - \log\omega(r,n) - n\log(1-s(1-R/r)) && (22)
\end{aligned}
$$

where (19) holds because G_s pushes forward μ_0 to μ_s and (21) holds by the Monge-Ampere formula for the density of μ_s with respect to Lebesgue measure and $\omega(r,n)$ is the volume of a ball of radius r in $\mathbb{R}^n$.

We wish to differentiate this expression with respect to s so we will need to justify differentiating under the integral sign. Firstly, note that $\phi(G_s(x))$ is integrable for all s because $\phi\circ G_s$ is continuous and μ_0 has compact support so $\phi\circ G_s$ is bounded μ_0-a.e. and μ_0 is a probability measure. Also $\frac{\partial(\phi\circ G_s)}{\partial s}$ is continuous since $\phi\in C^2$ and G_s is smooth in s. Furthermore, around any point s^* we wish to differentiate at, we may take a closed ball $\overline{B_\epsilon(s^*)}$. Then $\overline{B_\epsilon(s^*)}\times spt\mu_0$ is compact so $|\frac{\partial(\phi\circ G_s)}{\partial s}|$ achieves a maximum on it. We may then dominate using this constant bound which is again integrable because μ_0 is a probability measure. This shows that we may differentiate under the integral sign. Further, this also applies to the second derivative because $\phi \in C^2$. We then obtain the following formulas:

$$\frac{d}{ds}S(\mu_s) = \int_M (\nabla\phi(G_s(x)) \cdot \kappa) d\mu_0 - \frac{n(R/r-1)}{1-s(1-R/r)} \tag{23}$$

$$\frac{d^2}{ds^2}S(\mu_s) = \int_M (\kappa\nabla^2\phi(G_s(x))\kappa) d\mu_0 + \frac{n(R/r-1)^2}{(1-s(1-R/r))^2} \tag{24}$$

where $\kappa := (R/r-1)x + (x^*(1-R/r)+t)$. Let us investigate the convexity condition at $s = 0$. We need to show that the following inequality does not hold for some choice of r, R, x^*, t :

$$\int_M (\kappa\nabla^2\phi(x)\kappa) d\mu_0 + n(R/r-1)^2 - \frac{(\int_M (\nabla\phi(x)\cdot\kappa)d\mu_0 - n(R/r-1))^2}{N} \geq K d_2(\mu_0,\mu_1)^2 \tag{25}$$

$$\frac{\int_M (\kappa\nabla^2\phi(x)\kappa) d\mu_0 + n(R/r-1)^2}{N} - (\frac{\int_M (\nabla\phi(x)\cdot\kappa)d\mu_0 - n(R/r-1))}{N})^2 \geq \frac{K d_2(\mu_0,\mu_1)^2}{N} \tag{26}$$

$$\frac{\int_M (\kappa\nabla^2\phi(x)\kappa) d\mu_0}{N} - (\frac{\int_M (\nabla\phi(x)\cdot\kappa)d\mu_0}{N})^2 + \frac{2\int_M (\nabla\phi(x)\cdot\kappa)d\mu_0 * (R/r-1)}{N} \geq \frac{K d_2(\mu_0,\mu_1)^2}{N} \tag{27}$$

where (27) follows from setting $N = n$. When $n = N$ and ϕ is constant, the integrals vanish and the left hand side is zero. Therefore, convexity of the entropy will be violated if $K > 0$. When ϕ is not constant and K can be arbitrary, we proceed as follows. Since ϕ is not constant, $\nabla\phi$ is not identically zero and since it is continuous, the set where it is nonzero is nonempty open. Let x^* be such a point and set $t = -\nabla\phi(x^*)$ so that $\nabla\phi(x^*)\cdot\kappa(x^*) < 0$. We may rewrite $\kappa = (R/r-1)(x-x^*)+t$. Note that the first term is bounded in magnitude by $R - r$ $\mu_0(r)$-a.e so that $||\kappa - t|| < R$ $\mu_0(r)$-a.e. Since $\nabla\phi\cdot\kappa$ is continuous in x and κ, there exists a $\delta > 0$ so that if $r < \delta$ and $R < \delta$ then $\nabla\phi(x)\cdot\kappa \leq C < 0$ $\mu_0(r)$-a.e. So set $R = \delta/2$ and let $r \to 0$.

The first term of (27) is bounded as $r \to 0$. $\kappa\nabla^2\phi(x)\kappa$ is continuous in κ and x and these can both be taken to reside on constant compact subsets as shown above. The same argument works for the second term. As for the third term, by construction, its limsup goes to $-\infty$ as the integral is bounded above by a negative constant and $R/r - 1 \to \infty$ as $r \to 0$. Finally, the Wasserstein distance squared remains bounded as $r \to 0$ by the bound on distances between points in the balls:

$$\begin{aligned} d_2(\mu_0,\mu_1)^2 &= \inf_{\pi\in\Gamma(\mu_0,\mu_1)} |x-y|^2 d\pi(x,y) && (28)\\ &\leq \inf_{\pi\in\Gamma(\mu_0,\mu_1)} (|t|+r+R)^2 d\pi(x,y) && (29)\\ &= (|t|+r+R)^2 && (30) \end{aligned}$$

Therefore, for these choices of x^*, t, R and as $r \to 0$, we obtain a violation of convexity of the entropy for any K.

(c) We may recycle the previous part. Expanding (26), we obtain:

$$\begin{aligned} &\frac{\int_M (\kappa\nabla^2\phi(x)\kappa) d\mu_0}{N} - (\frac{\int_M (\nabla\phi(x)\cdot\kappa)d\mu_0}{N})^2 + \frac{2\int_M (\nabla\phi(x)\cdot\kappa)d\mu_0 * n(R/r-1)}{N^2} \\ &\qquad + \frac{n(R/r-1)^2}{N} - (\frac{n}{N})^2 (R/r-1)^2 \geq \frac{K d_2(\mu_0,\mu_1)^2}{N} \end{aligned} \tag{31}$$

Since $0 < N < n$, $(\frac{n}{N})^2 > \frac{n}{N}$. Thus if we take $r \to 0$, the negative last term will dominate the second last. If ϕ is constant, the first three terms vanish and the left hand side goes to $-\infty$. As we have seen, the right hand side is bounded. Therefore, convexity of the entropy fails in this case.

If ϕ is not constant, we may use the same procedure as in part (b). As before, the first two terms and the right hand side remain bounded as $r \to 0$. The final two terms on the left hand side sum to give a negative coefficient on $(R/r-1)^2$. The limsup of the third term goes to $-\infty$ as before. So the limsup of the left hand side again goes to $-\infty$ while the right hand side is bounded so convexity of the entropy fails here as well.

3. (a) We wish to show that for all $x, y \in M$, there exists a geodesic between them along which $V := V_1 + V_2$ is $((K_1 + K_2)d(x, y)^2, N_1 + N_2)$ convex. So let $x, y \in M$. Choose the geodesic between them such that V_1 is $(K_1 d(x, y)^2, N_1)$ convex along it, parametrized by γ. It is enough to show that $(V \circ \gamma)'' \geq \frac{1}{N_1+N_2}((V \circ \gamma)')^2 + (K_1 + K_2)d(x, y)^2$. But we know the corresponding inequalities for V_1 and V_2 separately. So by adding, we obtain the following:

$$(V_1 \circ \gamma)'' + (V_2 \circ \gamma)'' \geq \frac{1}{N_1}((V_1 \circ \gamma)')^2 + \frac{1}{N_2}((V_2 \circ \gamma)')^2 + (K_1 + K_2)d(x, y)^2 \tag{32}$$

$$(V \circ \gamma)'' \geq \frac{1}{N_1 + N_2}((V_1 \circ \gamma)' + (V_2 \circ \gamma)')^2 + (K_1 + K_2)d(x, y)^2 \tag{33}$$

$$(V \circ \gamma)'' \geq \frac{1}{N_1 + N_2}((V \circ \gamma)')^2 + (K_1 + K_2)d(x, y)^2 \tag{34}$$

where (33) follows from Young's inequality. Explicitly, we wish to show:

$$\frac{1}{N_1}a^2 + \frac{1}{N_2}b^2 \geq \frac{1}{N_1 + N_2}(a + b)^2 \tag{35}$$

$$a^2 + \frac{N_2}{N_1}a^2 + \frac{N_1}{N_2}b^2 + b^2 \geq a^2 + 2ab + b^2 \tag{36}$$

$$\frac{N_2}{N_1}a^2 + \frac{N_1}{N_2}b^2 \geq 2ab \tag{37}$$

where the last inequality (37) is true by Young's inequality. Therefore, V is $(K_1 + K_2, N_1 + N_2)$ convex.

(b) We wish to show that for all $\mu_0, \mu_1 \in P_2(M)$, there exists a d_2-geodesic connecting them with the relative entropy $S(\mu_s)$ is $((K_1 + K_2)d_2(\mu_0, \mu_1)^2, N_1 + N_2)$ convex along it.

$$S(\mu_s) = \int_M \frac{d\mu_s}{d(e^{-V}m)} log(\frac{d\mu_s}{d(e^{-V}m)})d(e^{-V}m) \tag{38}$$

$$= \int_M \frac{d\mu_s}{dm} log(e^V \frac{d\mu_s}{dm})dm \tag{39}$$

$$= \int_M log(\frac{d\mu_s}{dm}) + V d\mu_s \tag{40}$$

Now the first term is known to be weakly (K_1, N_1) convex since $(M, d, m) \in CD^e(K_1, N_1)$. Then by part (a), it's enough to show that $\int_M V d\mu_s$ is strongly (K_2, N_2) convex. But by Theorem 7.21 in Villani (2008), one may characterize geodesics μ_s in $P_2(M)$ as $(e_s)_\#\Pi$ where Π is a dynamical optimal coupling and e_s is the evaluation map at time s. Therefore, we have:

$$\int_M V d\mu_s = \int_\Gamma V(e_s(\gamma))d\Pi = \int_\Gamma (V \circ \gamma)(s)d\Pi \tag{41}$$

where Γ is the set of geodesics in M. But for a fixed geodesic γ, $(V \circ \gamma)$ is $(K_2 d_2(\mu_0, \mu_1)^2, N_2)$ convex since V was strongly (K_2, N_2) convex. We make the following assumptions so that we may differentiate under the integral sign: we assume $(V \circ \gamma) \in C^2$ and that $\gamma \to \max_s |(V \circ \gamma)'| \in L^1$ and $\gamma \to \max_s |(V \circ \gamma)''| \in L^1$. (We may assume for a fixed s, $(V \circ \gamma) \in L^1$ because if not, $S(\mu_s)$ is infinite and there is nothing to check.) Then we can compute:

$$\frac{d^2}{ds^2}(\int_\Gamma (V\circ\gamma)(s)d\Pi) = \int_\Gamma \frac{d^2}{ds^2}(V\circ\gamma)(s)d\Pi \tag{42}$$

$$\geq \int_\Gamma (\frac{1}{N_2}(\frac{d}{ds}(V\circ\gamma)(s))^2 + K_2 d_2(\mu_0,\mu_1)^2) \tag{43}$$

$$\geq \frac{1}{N_2}(\int_\Gamma \frac{d}{ds}(V\circ\gamma)(s))^2 + K_2 d_2(\mu_0,\mu_1)^2 \tag{44}$$

$$= \frac{1}{N_2}(\frac{d}{ds}\int_\Gamma (V\circ\gamma)(s))^2 + K_2 d_2(\mu_0,\mu_1)^2 \tag{45}$$

where (43) follows from the fact that $(V\circ\gamma)$ is $(K_2 d_2(\mu_0,\mu_1)^2, N_2)$ convex and (44) follows from Jensen's inequality. Therefore, $\int_M V d\mu_s$ is strongly (K_2, N_2) convex and the result follows.

4. (a) Since f_k has Γ-limit f_∞, for any $y \in M$, by condition (ii), $f_\infty(y) = \lim_{k\to\infty} f_k(y_k)$ for some sequence $y_k \to y$. But $\lim_{k\to\infty} f_k(y_k) \geq \liminf_{l\to\infty} \min f_{k_l}$ for every subsequence k_l so $f_\infty(y) \geq \liminf_{l\to\infty} \min f_{k_l}$. (The minimum exists because the functions are lower semicontinuous on a compact space).

Let $x_k \in argmin f_k$. Since the space is a compact metric space, some subsequence x_{k_l} converges to a point x_∞. Then by condition (i), $f_\infty(x_\infty) \leq \liminf_{l\to\infty} f_{k_l}(x_{k_l}) = \liminf_{l\to\infty} \min f_{k_l}$. If we combine this with the previous inequality we get that $x_\infty \in argmin f_\infty$.

(b) Let $M = \{a, b\}$ with the discrete topology. M is then a compact metric space. Let

$$f_k(x) = \begin{cases} 1/k & \text{if x = a} \\ e^{-k} & \text{if x = b} \end{cases} \tag{46}$$

Since the space is discrete, any function is continuous so this is lower semicontinuous. I claim f_k has Γ-limit $f_\infty = 0$. Condition (i) is satisfied because $0 = f_\infty(x_\infty) \leq f_k(x_k)$ for all k, x_k, x_∞. Condition (ii) is satisfied because for a given x, one can simply take the constant sequence $x_k = x$.

However, argmin $f_k = \{b\}$. So the sequence $x_k \in argmin f_k$ is the constant sequence $x_k = b$, so no subsequence can converge to a. Nevertheless $a \in argmin f_\infty$.

References

Agueh, Martial, and Guillaume Carlier. "BARYCENTERS IN THE WASSERSTEIN SPACE." SIAM Journal on Mathematical Analysis, vol. 43, no. 1-2, Society for Industrial and Applied Mathematics, 2011, pp. 904–24, https://doi.org/10.1137/100805741.

Villani, Cédric. Optimal transport: old and new. Accessed online at https://cedricvillani.org/sites/dev/files/old_images/2012/08/preprint-1.pdf.