Math 612: Single Cell Analysis	**2019W Term 1**

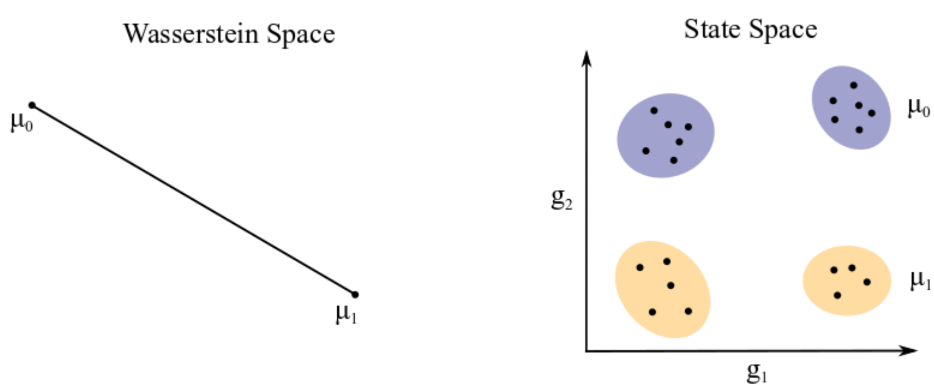
Lecture 15: October 24

Lecturer: Geoffrey Schiebinger | *Scribe: Sophie Boerlage*

15.1 Constant-speed geodesics in W_p

Theorem 15.1 *Geodesics from μ_0 to μ_1*

Let $\mu_0,\ \mu_1$ be two distributions in Wasserstein space, $W_p(\chi)$, where χ is convex state space, and $p > 1$.



Let Π^* be the optimal transport coupling of $\mu_0,\ \mu_1$

$$\Pi^* \leftarrow \operatorname*{argmin}_{\Pi} \mathbb{E}_{\Pi} \|x - y\|^P$$
$$X \sim \mu_0 \quad Y \sim \mu_1 \quad (X,\, Y) \sim \Pi$$

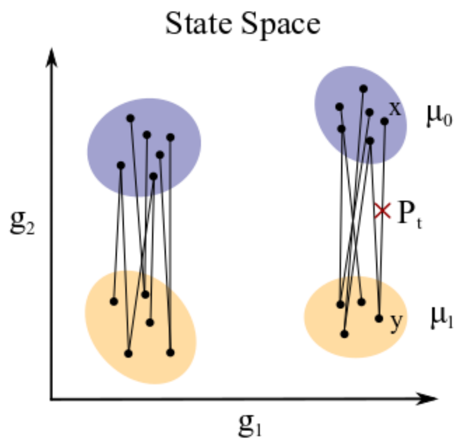
Let $P_t : \chi \times \chi \to \chi$

$$(x, y) \mapsto (1-t)x + ty$$

Def. The law of a random variable is the distribution it is sampled from.

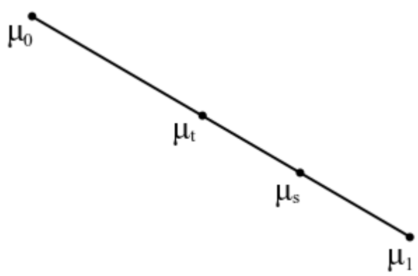
Let $\mu_t = Law(P_t(X, Y))$, where $X \sim \mu_0 \quad Y \sim \mu_1 \quad (X,\, Y) \sim \Pi^*$.

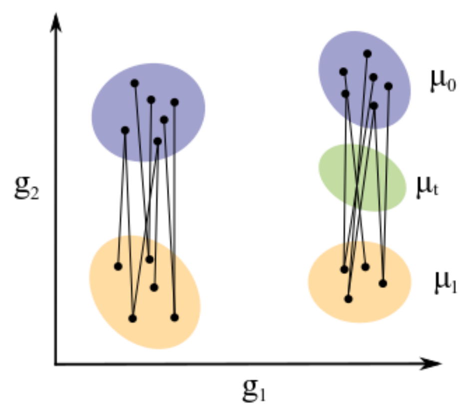
Sample an x, y pair and define the interpolated point P_t. Depending on the value of t, P_t will be closer to x or y.



Then, μ_t is a constant speed geodesic from μ_0 to μ_1:

$$W_p(\mu_t,\ \mu_s) = W_p(\mu_0,\ \mu_1)\,|t-s|$$





Proof: It suffices to prove

$$W_p(\mu_t,\ \mu_s) \le W_p(\mu_0,\ \mu_1)\,|t-s| \tag{15.1}$$

Indeed, assuming this, we have:

$$\begin{aligned} W_p(\mu_0,\ \mu_1) &\le W_p(\mu_0,\ \mu_t) + W_p(\mu_t,\ \mu_s) + W_p(\mu_s,\ \mu_t) \\ &\le t\,W_p(\mu_0,\ \mu_1) + (s-t)W_p(\mu_0,\ \mu_1) + (1-s)W_p(\mu_0,\ \mu_1) \\ &= W_p(\mu_0,\ \mu_1) \end{aligned}$$

The quantities we want are sandwiched in the equalities.

To prove 15.1, we consider a specific coupling of $\mu_t,\ \mu_s$. Let

$$\Pi_t^s = Law(P_t(X,Y),\ P_s(X,Y)) \quad (X,Y) \sim \Pi^*$$

Where $P_t(X, Y),\ P_s(X, Y)$ is a valid coupling of $\mu_t,\ \mu_s$. Let $P_t(X, Y) = z$, and $P_s(X, Y) = w$. Now, we have

$$\begin{aligned}
W_p(\mu_t,\ \mu_s) &\le (\mathbb{E}_{\Pi_t^s}\|z-w\|^p)^{1/p} \\
&= (\int \|z-w\|^p\, d\Pi_t^s(z,\ w))^{1/p} \\
&= (\int \|P_t(x,\ y) -\ P_s(x,\ y)\|^p\, d\Pi^*(x,\ y))^{1/p} \\
&= (\int \|(1-t)x + ty - (1-s)x - sy\|^p\, d\Pi^*(x,\ y))^{1/p} \\
&= (\int \|x - tx + ty - x + sx - sy\|^p\, d\Pi^*(x,\ y))^{1/p} \\
&= (\int \|(t-s)(x-y)\|^p\, d\Pi^*(x,\ y))^{1/p} \\
&= |t-s|(\int \|x-y\|^p\, d\Pi^*(x,\ y))^{1/p} \\
&= |t-s|\, W_p(\mu_0,\ \mu_1)
\end{aligned}$$

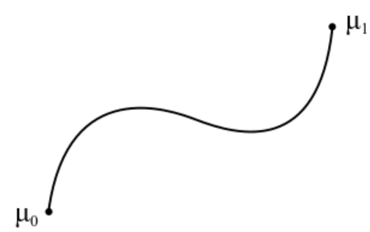
■

Reference: Proof can be found in Optimal Transport for Applied Mathematicians. The above theorem is 5.72, while theorem 5.14 is discussed below.

15.2 Curves in W_p

Theorem 15.2 *From curves to vector fields and back*

Let μ_t be a continuous curve in $W_p(\chi)$ for $p > 1$ and χ convex, $t \in (0, 1)$. Assume μ_t is a distribution with a density for (almost all) t.



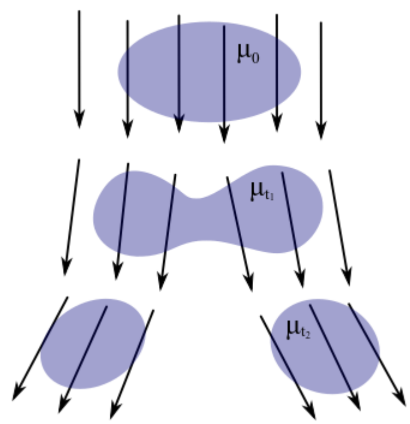
Then, there is a vector field v_t such that "u_t flows according to v_t."

$$\partial_t \mu_t + \nabla \cdot (v_t\, \mu_t) = 0$$

where $\nabla\cdot(v_t\,\mu_t)$ is the divergence or flux.

$$\begin{pmatrix}\partial_x\\\partial_y\\\partial_z\end{pmatrix}\cdot\begin{pmatrix}v_x\mu\\v_y\mu\\v_z\mu\end{pmatrix}=\partial v_x\,\mu+\partial v_y\,\mu+\partial v_z\,\mu$$

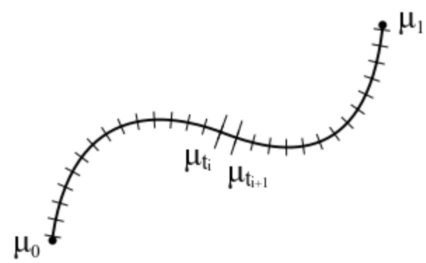
The reverse is also true. A vector field induces a curve.



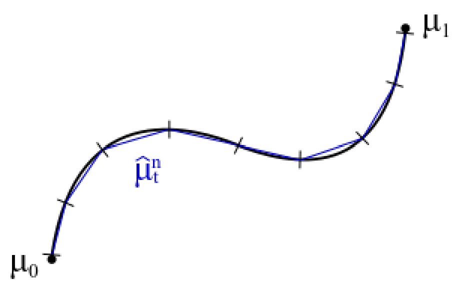
Note: The vector field is not necessarily the same as the developmental coupling of a developmental stochastic process. It does not account for entropy or growth.

Proof: $\mu_t \to v_t$

Divide (0, 1) into intervals of length $1/n$.



Consider the transport coupling Π_i^* connecting μ_{t_i} to $\mu_{t_{i+1}}$. Approximate the curve with geodesics $\hat{\mu}_t^n$ from μ_{t_i} to $\mu_{t_{i+1}}$.



Fact: For probability measures μ, ν on convex space χ with densities, the transport coupling $\Pi^* \leftarrow \mathrm{OT}(\mu, \nu)$ is deterministic. This means there is a function $f : \chi \to \chi$ that "does the transport."

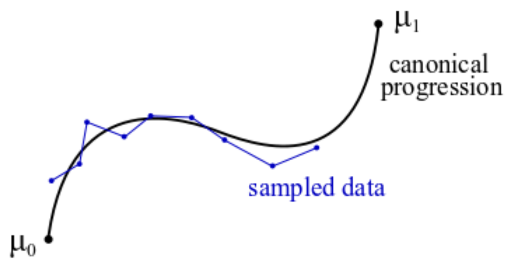
$$\Pi^*(x, \cdot\,) = f(x) \in \chi \quad \text{with probability 1}$$

$\Pi^*(x, \cdot\,)$ is a measure describing where x goes, and $f(x)$ is a point on state space because the function is deterministic.

Along each geodesic, $\hat{\mu}_t$ flows according to Π_i^* which can be described by a function f_i. To complete the proof, stitch the f_i together to get a piecewise constant vector field (in time). $\hat{\mu}_t$ flows according to this vector field.

■

While there is no growth or entropy, it can give a good approximation to the average or canonical path.



Fit the vector field to summarize couplings $\hat{\gamma}_{t_i t_{i+1}}$

Set up a regression to estimate a deterministic gene regulatory function, $X_{t_{i+1}} \simeq f(X_{t_i})$. Sample pairs $(X_{t_i}, X_{t_{i+1}}) \sim \hat{\gamma}_{t_i t_{i+1}}$

$$\min_{f \in \mathcal{F}} \mathbb{E}_{\hat{\gamma}_{t_i t_{i+1}}} \|X_{t_{i+1}} - f(X_{t_i})\|^2$$