Pathwise Free–Energy Minimization on the Probability Simplex

1 Model and Problem Statement

Consider a discrete probability simplex

$$\Delta^{m-1} = \Big\{ q \in \mathbb{R}^m_{\geq 0} : \mathbf{1}^\top q = 1 \Big\},$$

an initial point $q_0 \in \Delta^{m-1}$, and a preferred (target) distribution $p^\star \in \Delta^{m-1}$. Define the *free energy* (relative to $p^\star$) as the Kullback–Leibler divergence

$$\mathcal{F}(q) = D(q \| p^\star) = \sum_{i=1}^{m} q_i \log \frac{q_i}{p_i^\star}. \tag{1}$$

Each point $q \in \Delta^{m-1}$ may have a state–dependent cost $c(q) \geq 0$ that modulates the local contribution of free energy.

Let a discrete trajectory $\{q_k\}_{k=0}^T \subset \Delta^{m-1}$ connect the initial distribution q_0 to the terminal point q_T, and define the per–step cost

$$\ell(q_k, q_{k+1}) = c(q_k)\, \mathcal{F}(q_k) + \lambda\, D(q_{k+1} \| q_k), \qquad \lambda > 0. \tag{2}$$

The first term penalizes occupancy of high–free–energy regions (weighted by $c(\cdot)$), and the second term penalizes large changes in belief in the information–geometric sense. A terminal penalty $\phi(q_T) = \mathcal{F}(q_T)$ softly enforces proximity to $p^\star$.

The total cost of a path is then

$$J(\{q_k\}) = \sum_{k=0}^{T-1} \ell(q_k, q_{k+1}) + \phi(q_T). \tag{3}$$

The goal is to find the trajectory minimizing accumulated free energy:

$$\min_{q_1, \ldots, q_T \in \Delta^{m-1}} J(\{q_k\}). \tag{4}$$

Problem (4) defines a discrete geodesic on the simplex with respect to the KL geometry, subject to state–dependent energy weights $c(q)$.

2 Dynamic Programming Formulation

Let $V_k(q)$ denote the minimal remaining cost starting from q at stage k:

$$V_k(q) = \min_{q_{k+1}, \ldots, q_T} \left\{ \sum_{t=k}^{T-1} \Big[c(q_t)\, \mathcal{F}(q_t) + \lambda D(q_{t+1} \| q_t) \Big] + \phi(q_T) \right\}. \tag{5}$$

The terminal condition is

$$V_T(q) = \phi(q) = \mathcal{F}(q), \tag{6}$$

and the Bellman recursion for $k = T-1, \ldots, 0$ is

$$V_k(q) = c(q)\,\mathcal{F}(q) + \min_{q' \in \Delta^{m-1}} \Big\{ \lambda\, D(q' \| q) + V_{k+1}(q') \Big\}. \tag{7}$$

The minimizing argument

$$\pi_k(q) = \arg \min_{q' \in \Delta^{m-1}} \{ \lambda D(q' \| q) + V_{k+1}(q') \} \tag{8}$$

gives the optimal next point on the simplex. Equation (7) is a KL–proximal (mirror–descent) update, and the sequence $\{q_k\}$ generated by the policy (8) minimizes the accumulated, state–weighted free energy.

3 Worked Example on the 3–Simplex

We illustrate the Bellman method on a small discrete system with $m = 3$. The simplex is discretized to five representative points:

$$\begin{aligned} A &= (0.8,\, 0.1,\, 0.1), \\ B &= (0.6,\, 0.25,\, 0.15), \\ C &= (0.4,\, 0.3,\, 0.3), \\ D &= (0.3,\, 0.45,\, 0.25), \\ E &= (0.2,\, 0.5,\, 0.3) = p^\star. \end{aligned}$$

The free energies (1) relative to $p^\star = (0.2, 0.5, 0.3)$ are

$$\begin{aligned} &\mathcal{F}(A) = 0.8382, \quad \mathcal{F}(B) = 0.3819, \quad \mathcal{F}(C) = 0.1240, \\ &\mathcal{F}(D) = 0.0286, \quad \mathcal{F}(E) = 0. \end{aligned}$$

A state–dependent cost function penalizes region C:

$$c(A) = 1, \quad c(B) = 1, \quad c(C) = 6, \quad c(D) = 1, \quad c(E) = 1.$$

Allowed transitions are

$$A \to \{B, C\}, \quad B \to \{D, E\}, \quad C \to \{D, E\}, \quad D \to \{E\}, \quad E \to \{E\}.$$

With $\lambda = 1$, the KL transition costs are

$$\begin{aligned} &D(B\|A) = 0.1173, \quad D(C\|A) = 0.3819, \\ &D(D\|B) = 0.1843, \quad D(E\|B) = 0.3348, \\ &D(D\|C) = 0.0506, \quad D(E\|C) = 0.1168, \\ &D(E\|D) = 0.0263. \end{aligned}$$

The horizon is $T = 3$, and the terminal cost $\phi(q_T) = \mathcal{F}(q_T)$.

Value Iteration

Terminal stage.

$$V_3(q) = \mathcal{F}(q).$$

Stage $k = 2$. For each state, compute $V_2(q) = \min_{q'}[c(q)\mathcal{F}(q) + D(q'\|q) + V_3(q')]$. The minimal values are

$$\begin{aligned} V_2(A) &= 1.3374, & \pi_2(A) &= B, \\ V_2(B) &= 0.5948, & \pi_2(B) &= D, \\ V_2(C) &= 0.8233, & \pi_2(C) &= D, \\ V_2(D) &= 0.0549, & \pi_2(D) &= E, \\ V_2(E) &= 0, & \pi_2(E) &= E. \end{aligned}$$

Stage $k = 1$. Using the same recursion,

$$\begin{aligned} V_1(A) &= 1.5503, & \pi_1(A) &= B, \\ V_1(B) &= 0.6211, & \pi_1(B) &= D, \\ V_1(C) &= 0.8496, & \pi_1(C) &= D, \\ V_1(D) &= 0.0549, & \pi_1(D) &= E, \\ V_1(E) &= 0, & \pi_1(E) &= E. \end{aligned}$$

Stage $k = 0$. Finally,

$$\begin{aligned} V_0(A) &= \min_{q' \in \{B,C\}} \big\{c(A)\mathcal{F}(A) + D(q'\|A) + V_1(q')\big\} \\ &= \min\{0.8382 + 0.1173 + 0.6211,\ 0.8382 + 0.3819 + 0.8496\} \\ &= 1.5766, \end{aligned}$$

so that

$$V_0(A) = 1.5766, \qquad \pi_0(A) = B.$$

Optimal Path and Cost

Following the greedy policy sequence π_0, π_1, π_2, the optimal trajectory from A to E is

$$\boxed{A = (0.8, 0.1, 0.1) \ \to\ B = (0.6, 0.25, 0.15) \ \to\ D = (0.3, 0.45, 0.25) \ \to\ E = (0.2, 0.5, 0.3)}.$$

The total minimal cost is $V_0(A) = 1.5766$, composed of

$$\begin{aligned} A \to B: &\quad c(A)\mathcal{F}(A) + D(B\|A) = 0.9555, \\ B \to D: &\quad c(B)\mathcal{F}(B) + D(D\|B) = 0.5662, \\ D \to E: &\quad c(D)\mathcal{F}(D) + D(E\|D) = 0.0549. \end{aligned}$$

The path avoids the high–cost region C even though $\mathcal{F}(C)$ is small, illustrating the trade–off between *free–energy reduction* and *state–dependent penalties*. It also prefers two small KL steps ($B \to D \to E$) to a single large jump ($B \to E$), revealing the geodesic–like smoothing effect induced by the KL regularization.

4 Discussion

The Bellman recursion (7) provides a principled way to compute minimum–free–energy paths on the simplex. In the continuous–time limit, it yields a gradient flow in the Fisher–Rao metric with potential $c(q)\,\mathcal{F}(q)$, thus connecting information geometry, variational inference, and optimal control.