Dynamic Scaled Sampling for Deterministic Constraints Supplemental Material

Lei Li
University of California, Berkeley
leili@cs.berkeley.edu

Bharath Ramsundar
Stanford University
rbharath@stanford.edu

Stuart Russell
University of California, Berkeley
russell@cs.berkeley.edu

Technical details

Problem 3 (Posterior with bounded non-negative continuous variables) Let s be a target sum, ϵ be a positive number, and suppose there are k *independent* real-valued random variables $X_1, \dots, X_k$, such that $X_i \sim p_i(x)$ and the prior probability densities $p_i(x) > 0$ if and only if $x \in [0, \infty)$. ϵ is coded in finite bits. $p_i(\cdot)$ is represented in parametric form with finitely many parameters in constant bits. The goal is to find a sample (x_i) from the posterior probability density $p(X_1, \dots, X_k \mid \sum_{i=1}^k X_i = s)$ such that $\sum_{i=1}^k X_i = s$ and the un-normalized probability density

$$p(X_1 = x_1) \cdot p(X_2 = x_2) \cdots p(X_k = x_k) > \epsilon \tag{1}$$

If such an (x_i) does not exist, failure must be reported.

Theorem 2: Problem 3 is NP-hard.

Proof. Suppose that such a subset $A \subseteq \{a_1, \dots, a_k\}$ does exist. Let $X_i = x_i = a_i$ if $a_i \in A$ and $X_i = x_i = 0$ otherwise. Then by definition of $p_i(x_i)$:

$$\prod_{i=1}^{k} p(X_i = x_i)$$
$$= \frac{1}{\prod_i c_i} \cdot \frac{1}{(2\delta\sqrt{2\pi})^k} \prod_i \left(1 + \exp(-\frac{a_i^2}{\delta^2})\right) > \epsilon$$

By the definition of η and ϵ, the above LHS is greater than ϵ. As a result, the sample $(X_1 = x_1 \dots X_k = x_k)$ is a valid solution for Problem .

Conversely, if Problem has a valid solution, we can show there is a subset A of $\{a_1, \dots, a_k\}$ summing up to s. To show this, suppose that a sample $(X_1 = x_1, \dots, X_k = x_k)$ such that $\sum_{i=1}^k x_i = s$ and the un-normalized probability density $\prod_i p(X_i = x_i) > \epsilon$. We claim $|x_i| < \frac{1}{2k}$ or $|x_i - a_i| < \frac{1}{2k}$ for each $i = 1 \dots k$. To see this, note that if $|x_j| \geq \frac{1}{2k}$ and $|x_i - a_i| \geq \frac{1}{2k}$,

$$\prod_{i=1}^{k} p(X_i = x_i)$$
$$= \frac{1}{\prod_i c_i} \cdot \frac{1}{(2\delta\sqrt{2\pi})^k} \prod_i \left(\exp(-\frac{x_i^2}{\delta^2}) + \exp(-\frac{(a_i - x_i)^2}{\delta^2})\right)$$
$$\leq \frac{1}{\prod_i c_i} \cdot \frac{1}{(2\delta\sqrt{2\pi})^k} \prod_i \left(1 + \exp(-\frac{a_i^2}{\delta^2})\right)$$
$$\cdot \frac{\exp(-\frac{1}{4k^2\delta^2}) + \exp(-\frac{(a_j - 1/2k)^2}{\delta^2})}{1 + \exp(-\frac{a_i^2}{\delta^2})}$$
$$< \epsilon$$

Since

$$\frac{\exp(-\frac{1}{4k^2\delta^2}) + \exp(-\frac{(a_j - 1/2k)^2}{\delta^2})}{1 + \exp(-\frac{a_i^2}{\delta^2})} < 2\exp(-16) = \eta$$

We can select a set $A = \{a_i \| x_i - a_i| < \frac{1}{2k}\}$. It then follows $|\sum_a \in Aa - \sum_i x_i| < \frac{k}{2k} = \frac{1}{2}$. Therefore $\sum_{a \in A} a = s$. □

Lemma 1. *Δ_x^k is compact.*

Proof of Lem. 1

Proof. Δ_x^k is a closed and bounded subset of $\mathbb{R}^{k+1}$. □

Lemma 2. *Let f be a continuous real valued function defined on compact set X. Then f is bounded on X.*

Proof of Lem. 2

Proof. $f(X)$ is a compact subset of $\mathbb{R}$ and is hence closed and bounded. □

Lemma 3. *The target joint distribution $X_1, \dots, X_k, \sum_{i=1}^k X_i = S$ has continuous probability density function p on R^k with support containing the standard $k-1$ simplex of value S. That is*

$$\text{supp}(X_1, \dots, X_k, \sum_{i=1}^{k} X_i = S) \supseteq \Delta_S^{k-1}$$

Proof of Lem. 3

Proof. By Assumption (1) we can write the joint density as

$$N(x_1, \ldots, x_k, s) = \left(\prod_{i=1}^{n} p(x_i)\right) I\left(\sum_{i=1}^{k} x_i = s\right)$$

This expression is continuous on Δ_S^{k-1} when $s = S$ so we have a continuous density on Δ_S^{k-1}

□

Lemma 4. *If $S - \sum_{j=1}^{i-1} x_j > 0$ then*

$$P_{q_i(\cdot|\eta_i)}(0 \le X_i \le S - \sum_{j=1}^{i-1} x_j) \ge \frac{1}{2}$$

Proof of Lem. 4

Proof. Note that

$$\mathbb{E}_{q_i(\cdot|\eta_i)}[X_i] = \int x_i q_i(x_i|\eta_i) dx_i = \eta_i = \frac{S - \sum_{j=1}^{i-1} x_j}{k - i + 1}$$

Using this equation, we calculate that

$$\begin{aligned}
&P_{q_i(\cdot|\eta_i)}(0 \le X_i \le S - \sum_{j=1}^{i-1} x_j) \\
&\quad = \int q_i(x_i|\eta_i) I(0 \le x_i \le \eta_i(k - i + 1) dx_i \\
&\quad \ge 1 - \int q_i(x_i|\eta_i) I(x_i > 2\eta_i) \\
&\quad = 1 - P_{q_i(\cdot|\eta_i)}(x_i > 2\eta_i) \\
&\quad \ge 1 - \frac{\mathbb{E}_{q_i(\cdot|\eta_i)}[X_i]}{2\eta_i} = 1 - \frac{\eta_i}{2\eta_i} = \frac{1}{2}
\end{aligned}$$

□

Lemma 5. DYSC(S, k, X_i, q_i) *has nonzero probability density function q on $\mathbb{R}^k$ with support equal to the standard $k - 1$ simplex of value S*

$$\text{supp}(\text{DYSC}(S, k, X_i, q_i)) = \Delta_S^{k-1}$$

q is continuous in the interior of Δ_S^{k-1}, $\text{int}(\Delta_S^{k-1})$, and is lower bounded on Δ_S^{k-1} by some $c > 0$.

Proof of Lem. 5

Proof. We prove this result by constructing an explicit formula for q on $\mathbb{R}^k$. Given $x \in \mathbb{R}^k$, as before let

$$\ell(x, S) = \min(k - 1, \max_j \sum_{i=1}^{j-1} |x_i| < 1)$$

We need the function $\ell(x, S)$ since Alg. 1 stops random sampling at j if $j = k - 1$ or if $\sum_{i=1}^{j} x_j = S$. Using this function, we can write our formula,

$$q(x_1, \ldots, x_k) \propto \left(\prod_{i=1}^{\ell(x,S)} \frac{q_i(x_i|\eta_i) I\left(0 \le x_i \le S - \sum_{j=1}^{i-1} x_j\right)}{\int q_i(y|\eta_i) I\left(0 \le y \le S - \sum_{j=1}^{i-1} x_j\right) dy}\right)$$

If $x \notin \Delta_S^{k-1}$, the indicator functions in the above formula guarantee that $q(x) = 0$. If $x \in \Delta_S^{k-1}$, then the above formula simplifies.

$$\begin{aligned}
q(x_1, \ldots, x_i) &= \frac{N(x_1, \ldots, x_k)}{D(x_1, \ldots, x_k)} \\
N(x) &= \prod_{i=1}^{\ell(x,S)} q_i(x_i|\eta_i) \\
D(x) &= \prod_{i=1}^{\ell(x,S)} \int q_i(y|\eta_i) I(0 \le y \le S - \sum_{j=1}^{i-1} x_j) dy
\end{aligned}$$

where N and D are respectively the numerator and denominator. Assumption (3) guarantees that $N(x)$ and $D(x)$ are positive for $x \in \Delta_S^{k-1}$ and continuous in any neighborhood where $\ell(x, S)$ is constant. Specifically note that $\ell(x, S)$ is continuous in $\text{int}(\Delta_S^{k-1})$ so $N(x)$ and $D(x)$ are continuous there as well. Before we prove continuity of q on $\text{int}(\Delta_S^{k-1})$, we will need the following consequence of lemma 4

$$\begin{aligned}
&D(x_1, \ldots, x_k) \\
&\quad = \prod_{i=1}^{\ell(x,S)} \int q_i(y|\eta_i) I(0 \le y \le S - \sum_{j=1}^{i-1} x_j) dy \\
&\quad = \prod_{i=1}^{\ell(x,S)} P_{q_i(\cdot|\eta_i)}(0 \le X_i \le S - \sum_{j=1}^{i-1} x_j) \\
&\quad \ge \prod_{i=1}^{\ell(x,S)} \frac{1}{2} \quad \ge \quad \left(\frac{1}{2}\right)^k
\end{aligned}$$

Thus the denominator D is lower bounded by a positive number. Now let $x, y \in \text{int}(\Delta_S^{k-1})$. We can calculate

$$\begin{aligned}
|q(x) - q(y)| &= \left| \frac{N(x)}{D(x)} - \frac{N(y)}{D(y)} \right| \\
&= \left| \frac{N(x)}{D(x)} - \frac{N(x)}{D(y)} + \frac{N(x)}{D(y)} - \frac{N(y)}{D(y)} \right| \\
&= \left| N(x) \left(\frac{1}{D(x)} - \frac{1}{D(y)} \right) + \frac{1}{D(y)} \left(N(x) - N(y) \right) \right| \\
&\leq N(x) \left| \frac{1}{D(x)} - \frac{1}{D(y)} \right| + \frac{1}{D(y)} \left| N(x) - N(y) \right| \\
&\leq \left| \frac{1}{D(x)} - \frac{1}{D(y)} \right| + 2^k \left| N(x) - N(y) \right|
\end{aligned}$$

Since $\frac{1}{D}$ and N are both positive continuous functions on $\text{int}(\Delta_S^{k-1})$, we can make the right hand side arbitrarily small as $|x - y| \to 0$. It follows that q is continuous on $\text{int}(\Delta_S^{k-1})$. We now show that q is lower bounded on Δ_S^{k-1}. By Assumption (3), there exist continuous positive functions f_i on $[0, \infty)$ that lower bound q_i regardless of choice of η_i. Then $f = \prod_{i=1}^k f_i$ defines a positive continuous function on Δ_S^{k-1}. Since Δ_S^{k-1} is compact, f has a lower bound $c > 0$ on Δ_S^{k-1}. It follows that q is lower bounded by c. □