Supplementary material for AAAI 2021 'Computing an Efficient Exploration Basis for Learning with Univariate Polynomial Features'

by

CHAITANYA AMBALLA, MANU K. GUPTA AND SANJAY P. BHAT

The proof of Theorem 1 uses the following lemma.

Lemma 1. *Let $s \in [p_{\min}, p_{\max}]$ and suppose $p_1, \dots, p_{n+1} \in [p_{\min}, p_{\max}]$ are such that $p_i \neq p_j$ for all $i \neq j$. Then $c_1, \dots, c_{n+1} \in \mathbb{R}$ satisfy*

$$c_1 f_n(p_1) + \cdots + c_{n+1} f_n(p_{n+1}) = f_n(s) \tag{10}$$

if and only if $c_i = l_i(s, \mathbf{p})$ for each $i = 1, \dots, n+1$, where $\mathbf{p} = [p_1, \dots, p_{n+1}]^{\mathrm{T}}$, and $l_i(\cdot, \mathbf{p})$ is the ith Lagrange basis polynomial for the points $\{p_1, p_2, \dots, p_{n+1}\}$ given by

$$l_i(s, \mathbf{p}) \overset{\text{def}}{=} \frac{\prod_{j \neq i} (s - p_j)}{\prod_{j \neq i} (p_i - p_j)}. \tag{11}$$

Proof. Equation (10) may be rewritten as

$$V(\mathbf{p}) c(s) = f_n(s). \tag{12}$$

The determinant of the Vandermonde matrix $V(\mathbf{p})$ equals (see Fact 7.18.5 from Bernstein (2018))

$$\det(V(\mathbf{p})) = \prod_{1 \leq i < j \leq n+1} (p_j - p_i). \tag{13}$$

which is nonzero since $p_i \neq p_j$ for $j \neq i$. Equation (12) thus has a unique solution. Applying Cramer's rule (see Fact 3.16.12 from Bernstein (2018)) gives this solution to be

$$c_i = \frac{\det(V(\mathbf{p}_i^s))}{\det(V(\mathbf{p}))} \tag{14}$$

where $\mathbf{p}_i^s$ is the vector obtained by replacing the ith element of $\mathbf{p}$ by s. Using (13) to expand the determinants of the two Vandermonde matrices in (14) and canceling common terms gives $c_i = l_i(s, \mathbf{p})$. □

Proof of Theorem 1. To begin, we claim that the set D_n is not contained in a d-dimensional subspace of $\mathbb{R}^{n+1}$ for any $d < n+1$. Indeed, if $z \in \mathbb{R}^{n+1}$ is orthogonal to the subspace spanned by D_n, then $z^{\mathrm{T}} f_n(\cdot)$ is identically zero on $[p_{\min}, p_{\max}]$. Equating the first $n+1$ derivatives of the constant function $z^{\mathrm{T}} f_n(\cdot)$ to zero results in a triangular system of equations which is easily solved to obtain $z = 0$. Thus the orthogonal complement of the span of D_n is the trivial subspace $\{0\}$, and our claim follows.

To prove 1) implies 2), suppose the set $\{f_n(p_1), \dots, f_n(p_{n+1})\}$ is a barycentric spanner for D_n for some $\mathbf{p} = [p_1, \dots, p_{n+1}]^{\mathrm{T}} \in [p_{\min}, p_{\max}]^{n+1}$ such that $p_{\min} \leq p_1 \leq \cdots \leq p_{n+1} \leq p_{\max}$. Since the set D_n is not contained in any proper subspace of $\mathbb{R}^{n+1}$, the vectors $\{f_n(p_i)\}_{i=1}^{n+1}$ are all distinct. Hence, $p_i \neq p_j$ for $j \neq i$, and it follows that $p_{\min} \leq p_1 < \cdots < p_{n+1} \leq p_{\max}$.

Next, let $x = f_n(s)$ for some $s \in [p_{\min}, p_{\max}]$. By the definition of barycentric spanner, there exist $c(s) = [c_1(s), \dots, c_{n+1}(s)]^{\mathrm{T}} \in [-1, 1]^{n+1}$ such that $c_1(s) f_n(p_1) + \cdots + c_{n+1}(s) f_n(p_{n+1}) = f_n(s)$. By Lemma 1, $c_i(s) = l_i(s, \mathbf{p})$ for each i.

It now follows from the definition of a barycentric spanner that $|l_i(s, \mathbf{p})| \leq 1$ for all $s \in [p_{\min}, p_{\max}]$ and $i = 1, \dots, n+1$. On the other hand, directly substituting p_i in (11) gives $l_i(p_i, \mathbf{p}) = 1$ for all $i \in \{1, \dots, n+1\}$, showing that p_i is a local maximizer of $l_i(\cdot, \mathbf{p})$ for each i. We have already shown above that the points $p_2, \dots, p_n$ are necessarily interior points of the interval $[p_{\min}, p_{\max}]$. Hence first-order necessary conditions for optimality apply, and give

$$\left. \frac{\partial l_i(s, \mathbf{p})}{\partial s} \right|_{s = p_i} = 0, \ i = 2, \dots, n. \tag{15}$$

Using (11) in (15) directly yields (1). Next, note that

$$\left. \frac{\partial l_1(s, \mathbf{p})}{\partial s} \right|_{s = p_1} = \sum_{j \neq 1} \frac{1}{p_1 - p_j} < 0. \tag{16}$$

Hence, if $p_1 > p_{\min}$, then there exists $\epsilon > 0$ such that $p_1 - \epsilon \in [p_{\min}, p_{\max}]$ and $l_1(p_1 - \epsilon, \mathbf{p}) > l_1(p_1, \mathbf{p}) = 1$ which contradicts our earlier conclusion that $|l_1(s, \mathbf{p})| \leq 1$ for all $s \in [p_{\min}, p_{\max}]$. The contradiction shows that $p_1 = p_{\min}$. A similar argument shows that $p_{n+1} = p_{\max}$. This shows that 1) implies 2).

To show that 2) implies 3), consider a $\mathbf{p} \in \mathbb{R}^{n+1}$ as in the statement 2). On applying Proposition 1 with $k = n-1$, $a = p_1$, and $b = p_{n+1}$, we conclude that $z = [p_2, \dots, p_n]^{\mathrm{T}} \in \mathbb{R}^{n-1}$ is the unique global maximizer of the function U defined by (4). Comparing (13) with (4) shows that $\mathbf{p}$ is a maximizer of $\ln|\det(V(\cdot))|$ among all vectors $w \in \mathbb{R}^{n+1}$ satisfying $p_{\min} = w_1 < w_2 < \cdots < w_n < w_{n+1} = p_{\max}$. It follows that 2) implies 3).

To prove that 3) implies 1), suppose $\mathbf{p}$ is as in statement 3), and consider $s \in (p_{\min}, p_{\max})$. Arguing as in the proof of "1) implies 2)", we see that $c(s) \in \mathbb{R}^{n+1}$ defined by (14) satisfies $f_n(s) = c_1(s) f_n(p_1) + \cdots + c_n(s) f_n(p_{n+1})$. By the global optimality of $\mathbf{p}$, we have $|\det(V(\mathbf{p}_i^s))| \leq |\det(V(\mathbf{p}))|$, that is, $|c_i(s)| \leq 1$ for all i. This completes the proof. □

Proof of Proposition 1. First, observe that $C = \{z \in \mathbb{R}^k : a < z_1 < z_2 < \cdots < z_k < b\}$ is an open convex set. Note that the function $x \mapsto \ln|x - r|$ is continuously differentiable at $x \neq r$ with derivative $(x - r)^{-1}$. Using this observation, one can conclude that U is continuously differentiable on C, and calculate

$$\frac{\partial U}{\partial z_i}(z) = \frac{1}{z_i - a} + \sum_{j \neq i} \frac{1}{z_i - z_j} + \frac{1}{z_i - b}, \ i = 1, \dots, k. \tag{17}$$

We can differentiate (17), and further calculate

$$\frac{\partial^2 U}{\partial z_i^2} = \frac{-1}{(z_i - a)^2} - \sum_{j \neq i} \frac{1}{(z_i - z_j)^2} - \frac{1}{(z_i - b)^2}, \tag{18}$$

$$\frac{\partial^2 U}{\partial z_i z_j} = \frac{1}{(z_i - z_j)^2}, \tag{19}$$

for $i, j \in \{1, \dots, k\}$, $j \neq i$. The second-order mixed partial derivatives in (18) and (19) define the Hessian matrix $H(z)$ of U at $z \in C$. Applying the Gershgorin circle theorem (see Fact 6.10.22 from Bernstein (2018)) to $H(z)$ lets us conclude that $H(z)$ is negative definite for each $z \in C$. This implies that U is strictly concave on C.

We first show that U has a unique global maximizer in C. To show this, note that the function U is unbounded below.

For instance, $U \to -\infty$ as $z_1 \to a$. Hence we may choose $K \in \mathbb{R}$ such that the set $F \stackrel{\text{def}}{=} \{x \in C : U(x) \geq K\}$ is nonempty. We claim that F is closed in $\mathbb{R}^k$. To arrive at a contradiction, suppose F is not closed. Then there exists $x \in \mathbb{R}^k \backslash F$ and a sequence $\{x_l\}_{l=1}^{\infty}$ in F converging to x. Since $F \subseteq C$, x belongs to the closure of C. On the other hand, $x \notin C$, since otherwise the continuity of U on C would imply that $K \geq U(x_l) \to U(x)$, and contradict our assumption that $x \notin F$. Thus x lies in the closure of C, but not in C. It follows that x satisfies at least one of the inequalities defining C with equality. However, the definition of U then implies that the sequence $\{U(x_l)\}_{l=1}^{\infty}$ diverges to $-\infty$, contradicting our definition of F. This proves our claim that F is closed.

F is also bounded, and hence compact, as C itself is contained in the bounded set $[a, b]^k$. The continuous function U achieves its maximum over the compact set F at a point, say $z^* \in F$. By the definition of F, we have $U(z^*) \geq K$, while $U(z) < K \leq U(z^*)$ for all $z \in C \backslash F$. Thus we conclude that z^* is a global maximizer of U on C. Being strictly concave, U can have at most one global maximizer (Boyd and Vandenberghe 2004). It follows that z^* is the unique global maximizer of U on C.

Since C is open, first-order necessary conditions for optimality imply that the first-order partial derivatives of U given by (17) vanish at z^*. Thus, z^* is a solution to (3).

If $x \in C$ is any solution of (3), then, by (17), the gradient of U at x is zero, while the Hessian $H(x)$ is negative definite. By second-order sufficient conditions for optimality, x is a local maximizer for U. However, strict concavity implies that x is also a global maximizer of U on C. It now follows from the uniqueness of the global maximizer shown above that $x = z^*$. Thus z^* is the unique solution to (3).

Next, consider the point $x \in \mathbb{R}^k$ defined by setting $x_i = b + a - z^*_{k+i-1}$. It is a simple matter to check that $x \in C$, and verify by direct substitution that x satisfies (3). Since we have already shown that z^* is the unique solution to (3) in C, it follows that $x = z^*$. In other words, (5) holds. This completes the proof.

□

Reduced form of Equations (1) and (2) by exploiting symmetry

The relations (5) imply that the points p_i, $i = 1, \ldots, n+1$, yielding the barycentric spanner are symmetrically placed about the midpoint $\bar{p} \stackrel{\text{def}}{=} \frac{1}{2}(p_{\min} + p_{\max})$ of the interval $[p_{\min}, p_{\max}]$. Thus, it is sufficient to find points lying only on one side of the midpoint. This can be essentially achieved by using the symmetry relations (5) to eliminate (roughly) half the variables from (1) and (2). Next, we describe the reduced versions of (1) and (2) obtained by exploiting the symmetry inherent in (5).

First assume $n = 2l$ for some $l > 0$. Then $p_{l+1} = \bar{p}$ by symmetry, and solving (1) reduces to solving

$$\left[\sum_{\substack{j \neq i \\ 1 \leq j \leq l}} \frac{1}{p_i - p_j} + \frac{1}{p_i + p_j - 2\bar{p}}\right] + \frac{3}{2(p_i - \bar{p})} = 0, \quad (20)$$

for $i = 2, \ldots, l$. Likewise, optimizing (2) in the case $n = 2l$ reduces to optimizing the function

$$\bar{U}(p) = \ln \left| \left(\prod_{i=2}^{l} (a - p_i)^2 (b - p_i)^2 (p_i - \bar{p})^3 \right) \times \left(\prod_{2 \leq i < j \leq l} (p_i - p_j)^2 (p_i + p_j - 2\bar{p})^2 \right) \right|, \quad (21)$$

on the set $p_{\min} < p_1 < \ldots < p_l < \bar{p}$. Next, assume $n = 2l + 1$ for some $l > 0$. In this case, a solution of (1) can be recovered by solving

$$\sum_{\substack{j \neq i \\ 1 \leq j \leq l+1}} \frac{1}{p_i - p_j} + \frac{1}{p_i + p_j - 2\bar{p}} + \frac{1}{2(p_i - \bar{p})} = 0, \; i = 2, \ldots, l+1, \quad (22)$$

while the optimizer in (2) can be found by optimizing

$$\bar{U}(p) = \ln \left| \left(\prod_{i=2}^{l+1} (a - p_i)^2 (b - p_i)^2 (p_i - \bar{p}) \right) \times \left(\prod_{2 \leq i < j \leq l+1} (p_i - p_j)^2 (p_i + p_j - 2\bar{p})^2 \right) \right|, \quad (23)$$

on the set $p_{\min} < p_1 < \ldots < p_l < \bar{p}$.

Proof of Proposition 2

In order to prove Proposition 2, we first prove the following result.

Proposition 3. *A barycentric spanner for the set D solves the following minmax problem.*

$$\min_{x_1, \ldots, x_d \in D} \max_{z \in D} \|X^{-1} z\|_{\infty}. \quad (24)$$

Proof. Given a subset $\{x_1, \ldots, x_d\}$ of D and $X = [x_1, \ldots, x_d] \in \mathbb{R}^{d \times d}$, letting $z = x_1$ gives $\|X^{-1} z\|_{\infty}^2 = \|e_1\|_2^2 = 1$. Thus $\max_{z \in D} \|X^{-1} z\|_{\infty}^2 \geq 1$ for all choices of X. On the other hand, if $\{x_1, \ldots, x_d\}$ is a barycentric spanner for D, then $\|X^{-1} z\|_{\infty}^2 \leq 1$ for all $z \in D$. This proves that a barycentric spanner solves (24). □

Proof of Proposition 2: The expected mean-square testing error on the test points is

$$\frac{1}{k} \sum_{i=1}^{k} \mathbb{E}[\hat{g}(z_i) - g(z_i)]^2 = \frac{\sigma^2}{k} \operatorname{tr}(Z^{\mathrm{T}} (X X^{\mathrm{T}})^{-1} Z), \quad (25)$$

where $Z \stackrel{\text{def}}{=} [z_1, \ldots, z_k] \in \mathbb{R}^{d \times k}$. The learner's goal is to choose X such that the worst case value of the expected mean-square testing error in (25) over the adversary's choice of Z is minimized.

Let $\Lambda_1 \in \mathbb{R}^{d\times d}$ denote the diagonal matrix having σ_i as its ith diagonal entry for each i. Note that if $\epsilon = [\epsilon_1, \ldots, \epsilon_d]$, then $\mathbb{E}(\epsilon\epsilon^{\mathrm{T}}) = \Lambda_1^2$. Using this along with (7) gives

$$\begin{aligned}
&\frac{1}{k}\sum_{i=1}^{k}\mathbb{E}[\hat{g}(z_i) - g(z_i)]^2 \\
&= \frac{1}{k}\mathrm{tr}[Z^{\mathrm{T}}(XX^{\mathrm{T}})^{-1}X\Lambda_1^2X^{\mathrm{T}}(XX^{\mathrm{T}})^{-1}Z] \\
&= \frac{1}{k}\mathrm{tr}(Z^{\mathrm{T}}X^{-\mathrm{T}}\Lambda_1^2X^{-1}Z) = \frac{1}{k}\sum_{j=1}^{k}\|\Lambda_1X^{-1}z_j\|_2^2 \\
&= \frac{1}{k}\sum_{j=1}^{k}[\sigma_1^2(e_1^{\mathrm{T}}X^{-1}z_j)^2 + \cdots \\
&\quad + \sigma_d^2(e_d^{\mathrm{T}}X^{-1}z_j)^2]. \qquad (26)
\end{aligned}$$

It is easy to see from (26) that the adversary can ensure the worst case expected mean-square error for a given choice of X by setting $k = 1$, computing $(i^*, z^*) = \mathrm{argmax}_{i,z}|e_i^{\mathrm{T}}X^{-1}z|$, and setting $z_1 = z^*$, $\sigma_{i^*} = \sigma$ and $\sigma_i = 0$ for all $i \neq i^*$. Note that by definition $|e_{i^*}^{\mathrm{T}}X^{-1}z_*| = \max_{z\in D}\|X^{-1}z\|_\infty$. It is now evident from Proposition 3 that the learner can minimize the worst case expected mean-square error forced by the adversary by choosing the training points to form a *barycentric spanner* for the set D. □