Supplemental Notes for Bernoulli Factory Algorithms

This version of the document is dated 2021-02-07.

1 General Factory Functions

The algorithms for **general factory functions** work with two sequences of polynomials: one approaches the function $f(\lambda)$ from above, the other from below, where f is a continuous function that maps the interval (0, 1) to (0, 1). (These two sequences form a so-called *approximation scheme* for f.) One requirement for these algorithms to work correctly is called the *consistency requirement*:

- For each sequence, the difference between one polynomial and the previous one must have non-negative Bernstein coefficients (once the latter polynomial is elevated to the same degree as the other).

The consistency requirement ensures that the polynomials approach the target function without crossing each other. Unfortunately, the reverse is not true in general; even if the upper polynomials "decrease" and the lower polynomials "increase" to f, this does not mean that the scheme will ensure consistency.

1.1 Schemes That Don't Work

In the academic literature (papers and books), there are many approximation schemes that involve polynomials that converge from above and below to a function. Unfortunately, most of them cannot be used as is to simulate a function f in the Bernoulli Factory setting, because they don't ensure the consistency requirement described earlier.

The following are approximation schemes with counterexamples to consistency.

In this section, **fbelow** and **fabove** mean the k^{th} Bernstein coefficient for the lower or upper degree-n polynomial, respectively, where k is an integer in the interval [0, n].

First scheme. In this scheme (Powell 1981)[(1)], let f be a twice differentiable function (that is, a C^2 continuous function, or a function with continuous "slope" and "slope-of-slope" functions). Then the upper polynomial of degree n has Bernstein coefficients as follows, for all $n \geq 1$:

- **fabove**(n, k) = $f(k/n) + M / (8*n)$,

where M is an upper bound of the maximum absolute value of f's slope-of-slope function (second derivative), and where k is an integer in the interval [0, n].

And the lower polynomial of degree n has Bernstein coefficients as follows:

- **fbelow**(n, k) = f(k/n) + M / (8*n).

The counterexample involves the twice differentiable function $g(\lambda) = \sin(\pi*\lambda)/4 + 1/2$.

For g, the Bernstein coefficients for—

- the degree-2 upper polynomial (b(5, k)) are [0.6542..., 0.9042..., 0.6542...], and
- the degree-4 upper polynomial (b(6, k)) are [0.5771..., 0.7538..., 0.8271..., 0.7538..., 0.5771...].

The degree-2 polynomial lies above the degree-4 polynomial everywhere in [0, 1]. However, to ensure consistency, the degree-2 polynomial, once elevated to degree 4, must have Bernstein coefficients that are greater than or equal to those of the degree-4 polynomial.

- Once elevated to degree 4, the degree-2 polynomial's coefficients are [0.6542..., 0.7792..., 0.8208..., 0.7792..., 0.6542...].

As we can see, the elevated polynomial's coefficient 0.8208... is less than the corresponding coefficient 0.8271... for the degree-4 polynomial.

Second scheme. In this scheme, let *f* be a Lipschitz continuous function in [0, 1] (that is, a continuous function whose slope does not tend to a vertical slope anywhere in [0, 1]). Then the upper polynomial of degree *n* has Bernstein coefficients as follows, for all n≥2:

- **fabove**(n, k) = f(k/n) + (5/4) / sqrt(n),

where L is the maximum absolute "slope", also known as the Lipschitz constant, and (5/4) is the so-called Popoviciu constant, and where *k* is an integer in the interval [0, *n*] (Lorentz 1986)[(2)], (Popoviciu 1935)[(3)].

And the lower polynomial of degree *n* has Bernstein coefficients as follows, for all n≥1:

- **fbelow**(n, k) = f(k/n) + (5/4) / sqrt(n).

The following counterexamples show that this scheme can fail to ensure consistency, even if the set of functions is restricted to "smooth" functions (not just Lipschitz continuous functions).

For the first counterexample, the function $g(\lambda)$ = min(λ, 1−λ)/2 is Lipschitz continuous with Lipschitz constant 1. (In addition, *g* has a kink at 1/2, so that it's not differentiable, but this is not essential for the counterexample.)

For *g*, the Bernstein coefficients for—

- the degree-5 upper polynomial (b(5, k)) are [0.4874..., 0.5874..., 0.6874..., 0.6874..., 0.5874..., 0.4874...], and
- the degree-6 upper polynomial (b(6, k)) are [0.4449..., 0.5283..., 0.6116..., 0.6949..., 0.6116..., 0.5283..., 0.4449...].

The degree-5 polynomial lies above the degree-6 polynomial everywhere in [0, 1]. However, to ensure consistency, the degree-5 polynomial, once elevated to degree 6, must have Bernstein coefficients that are greater than or equal to those of the degree-6 polynomial.

- Once elevated to degree 6, the degree-5 polynomial's coefficients are [0.4874..., 0.5707..., 0.6541..., 0.6874..., 0.6541..., 0.5707..., 0.4874...].

As we can see, the elevated polynomial's coefficient 0.6874... is less than the corresponding coefficient 0.6949... for the degree-6 polynomial.

There is a similar counterexample that can be built:

- When *g* = sin(4*π*λ)/4 + 1/2, a "smooth" function with Lipschitz constant π, the counterexample is present between the degree-3 and degree-4 lower polynomials.

Thus, we have shown that this approximation scheme is not guaranteed to meet the consistency requirement for all Lipschitz continuous functions.

It is yet to be seen whether a counterexample exists for this scheme when *n* is restricted to powers of 2.

Third scheme. Same as the second scheme, but replacing (5/4) with the Sikkema constant, *S* = (4306+837*sqrt(6))/5832 (Lorentz 1986)**(2)**, (Sikkema 1961)**(4)**. In fact, the same counterexamples for the second scheme apply to this one, since this scheme merely multiplies the offset to bring the approximating polynomials closer to *f*.

For example, the first counterexample for this scheme is almost the same as the first one for the second scheme, except the coefficients for—

- the degree-5 upper polynomial are [0.5590..., 0.6590..., 0.7590..., 0.7590..., 0.6590..., 0.5590...], and
- the degree-6 upper polynomial are [0.5103..., 0.5936..., 0.6770..., 0.7603..., 0.6770..., 0.5936..., 0.5103...].

And once elevated to degree 6, the degree-5 polynomial's coefficients are [0.5590..., 0.6423..., 0.7257..., 0.7590..., 0.7257..., 0.6423..., 0.5590...].

As we can see, the elevated polynomial's coefficient 0.7590... is less than the corresponding coefficient 0.7603... for the degree-6 polynomial.

1.2 Other Schemes

I have found how to extend the results of Nacu and Peres (2005)**(5)** to certain functions with a slope that tends to a vertical slope. Moreover, the polynomials satisfy the consistency requirement unlike with the schemes from the previous section.

For example, take a factory function *f*(λ), the function to simulate using flips of a coin with unknown probability of heads of λ. The following scheme to build upper and lower polynomials can be used if *f*(λ)—

- is (1/2)-*Hölder continuous*, meaning its vertical slopes, if any, are no "steeper" than *m**sqrt(λ), for some number *m* greater than 0, and
- in the interval [0, 1]— - has a minimum of greater than 0 and a maximum of less than 1, or - is *convex* (the rate of growth of its "slope" never decreases) and has a minimum of greater than 0, or - is *concave* (the rate of growth of its "slope" never increases) and has a maximum of less than 1.

Finding the *Hölder constant m* is non-trivial and it requires knowing whether *f* has a vertical slope and where, among other things.**(6)** But assuming *m* is known, then for all *n* that are powers of 2:

- **fbelow**(n, k) = $f(k/n) - (m*(2^{1/4} + 2^{2/4} + 2^{3/4} + 1))/n^{1/4}$ (or $f(k/n)$ if *f* is concave).
- **fabove**(n, k) = $f(k/n) + (m*(2^{1/4} + 2^{2/4} + 2^{3/4} + 1))/n^{1/4}$ (or $f(k/n)$ if *f* is convex).

Proofs are in the appendix.

2 Notes

- [(1)] Powell, M.J.D., *Approximation Theory and Methods*, 1981
- [(2)] G. G. Lorentz. Bernstein polynomials. 1986.
- [(3)] Popoviciu, T., "Sur l'approximation des fonctions convexes d'ordre supérieur", Mathematica (Cluj), 1935.
- [(4)] Sikkema, P.C., "Der Wert einiger Konstanten in der Theorie der Approximation mit Bernstein-Polynomen", Numer. Math. 3 (1961).
- [(5)] Nacu, Şerban, and Yuval Peres. "**Fast simulation of new coins from old**", The Annals of Applied Probability 15, no. 1A (2005): 93-115.

- [(6)] Specifically, the constant *m* is an upper bound of abs($f(x)-f(y)$)/sqrt(abs($x-y$)) for all x, y pairs, where x and y are each in [0, 1] and x != y. However, this bound can't directly be calculated as it would involve checking an infinite number of x, y pairs.

3 Appendix

3.1 Proofs for Hölder Function Approximation Scheme

There is an easy extension to lemma 6(i) of Nacu and Peres (2005)**[(5)]** to certain functions with a slope that tends to a vertical slope. Specifically, it applies to any *Hölder continuous* function, which means a continuous function whose slope doesn't go exponentially fast to a vertical slope.

The parameters α and M, in the lemma below, mean that the function is no "steeper" than $M*\lambda^{\alpha}$; α is in the interval (0, 1] and M is greater than 0.

Lemma 1. _Let f(λ) be a continuous function that maps [0, 1] to [−1, 1], and let X be a hypergeometric(2*n, k, n) random variable. If f is α-Hölder continuous with Hölder constant M, then—

abs($\mathbf{E}[f(X/n)] - f(k/(2*n))$),   (1)

is bounded from above by $M*(1/(2*n))^{\alpha/2}$.

Proof. abs($\mathbf{E}[f(X/n)] - f(\text{k}/(2*n))$) ≤ $\mathbf{E}$[abs($f(X/n) - f(k/(2*n))$)] ≤ $M*\mathbf{E}[\text{abs}(X/n - k/(2*n))]^{\alpha}$ (by the definition of Hölder continuous functions) ≤ $M*(\mathbf{E}[\text{abs}(X/n - k/(2*n))]^2)^{\alpha/2}$ = $M*\mathbf{Var}[X/n]^{\alpha/2}$ ≤ $M*(1/(2*n))^{\alpha/2}$. □

Notes:

1. **E**[.] means expected or average value, and **Var**[.] means variance.
2. A *Lipschitz-continuous* function has no slope that tends to a vertical slope, making it a 1-Hölder continuous function with M equal to its Lipschitz constant.
3. An α-Hölder continuous function in [0, 1] is also β-Hölder continuous for any β less than α.

In fact, a tighter bound can be achieved as follows:

Lemma 2. *With the assumptions in Lemma 1, the expression (1) is bounded from above by* $M*(1/(7*n))^{\alpha/2}$*, for all integers* $n\ge4$.

Proof. For all integers $n\ge4$, abs($\mathbf{E}[f(X/n)] - f(\text{k}/(2*n))$) ≤ $M*\mathbf{Var}[X/n]^{\alpha/2}$ = $M*(k*$

$(2*n-k)/(4*(2*n-1)*n^2))^{\alpha/2} \le M*(n^2/(4*(2*n-1)*n^2))^{\alpha/2} = M*(1/(8*n-4)^{\alpha/2} \le M*(1/(7*n)^{\alpha/2}$.

Theorem 1. *Let f(λ) be an α-Hölder continuous function with Hölder constant M that maps [0, 1] to (0, 1). The following Bernstein coefficients (**fabove**(n, k) for the upper polynomials, and **fbelow**(n, k) for the lower polynomials) form an approximation scheme that meets conditions (i), (iii), and (iv) of Proposition 3 of Nacu and Peres (2005)[(5)], for all n≥1, and thus can be used to simulate f via the algorithms for general factory functions described at the top of this page:*

- ***fbelow**(n, k) = f(k/n) − δ(n) (kth Bernstein coefficient of lower nth degree polynomial).*
- ***fabove**(n, k) = f(k/n) + δ(n) (kth Bernstein coefficient of upper nth degree polynomial).*

Where δ(n) is a solution to either of the following functional equations:

- $\delta(n) = \delta(2*n) + M*(1/(2*n))^{\alpha/2}$.
- $\eta(n) = \eta(2*n) + M*(1/(7*n))^{\alpha/2}$.

Proof. Follows from Lemma 1 above as well as the proof of Proposition 10 of Nacu and Peres (2005)[(5)].

Theorem 2. *Let f be as given in Theorem 1. The following Bernstein coefficients form an approximation scheme that meets conditions (i), (iii), and (iv) of Proposition 3 of Nacu and Peres (2005)[(5)], for all n≥1:*

If n is 4 or greater:

- ***fbelow**(n, k) = f(k/n) − η(n).*
- ***fabove**(n, k) = f(k/n) + η(n).*

Otherwise:

- ***fbelow**(n, k) = Lower bound of **fbelow**(4, k) for all k in [0, 4].*
- ***fabove**(n, k) = Upper bound of **fabove**(4, k) for all k in [0, 4].*

Where η(n) is a solution to the functional equation $\eta(n) = \eta(2*n) + M*(1/(7*n))^{\alpha/2}$.

Proof. Follows from Lemma 2 above as well as the proof of Proposition 10 of Nacu and Peres, as well as from the observation in Remark B of the paper that we can start the algorithm from n = 4; in that case, the upper and lower polynomials of degree 1 through 3 above would be constant functions whose Bernstein coefficients are all the same. □

Proposition 1.

1. *Let f be as given in Theorem 1, except f is concave and may have a minimum of 0. The approximation scheme of Theorem 1 or 2 remains valid if **fbelow**(n, k) = f(k/n), rather than as given in either theorem.*
2. *Let f be as given in Theorem 1, except f is convex and may have a maximum of 1. The approximation scheme of Theorem 1 or 2 remains valid if **fabove**(n, k) = f(k/n), rather than as given in either theorem.*

Proof. Follows from Theorem 1 or 2 and Jensen's inequality. □

Unfortunately, there is no easy way to solve the functional equations above in a way that works for all α. However, the following examples show solutions that lead to approximation schemes that work for any α-Hölder continuous function with certain values of α.

- If α is 1/2 or greater: $\delta(n)$ = (M*($2^{1/4}$ + $2^{2/4}$ + $2^{3/4}$ + 1))/$n^{1/4}$. (Solved via SymPy: `rsolve(Eq(f(n),f(n+1)+z*(1/(2*2**n))**((S(1)/2)/2)),f(n)).subs(n,log(n,2)).simplify()`.)
- If α is 2/3 or greater: $\delta(n)$ = $2^{2/3}$*M*($2^{1/3}$ + $2^{2/3}$ + 2)/(2*$n^{1/3}$).
- If α is 1/j or greater: $\delta(n)$ = (M* $\sum_{i\,=\,0,\ldots,(j*2)-1}$ $2^{i/(2*j)}$)/$n^{1/(2*j)}$.
- If α is 1 (f is Lipschitz continuous): $\eta(n)$ = (M*sqrt(7)*(sqrt(2)+2)/(7*sqrt(n)). (Solved via SymPy: `rsolve(Eq(f(n),f(n+1)+z*(1/(7*2**n))**((S(1))/2)),f(n)).subs(n,log(n,2)).simplify()`.)
- If α is 1/2 or greater: $\eta(n)$ = (M*$7^{3/4}$*($2^{1/4}$ + $2^{2/4}$ + $2^{3/4}$ + 2))/(7*$n^{1/4}$).

4 License