Bayesian Integrals on Toric Varieties

Michael Borinsky, Anna-Laura Sattelberger, Bernd Sturmfels, Simon Telen

Hypergeometric School 2023
Kobe University

August 18, 2023
What to expect

Aim

Computation of marginal likelihood integrals

$$\int_{X > 0} p_0^{u_0} p_1^{u_1} \cdots p_m^{u_m} \Omega_{\text{prior}}^X$$

for statistical models that are parameterized by a toric variety.

How?

Tropical sampling algorithms.

Outline

1. Toric varieties and statistical models
2. Toric varieties as probability spaces
3. Tropical sampling

Toric varieties and statistical models

**Definition.**
A **discrete statistical model** taking $m + 1$ states is a parameterized subset of the probability $m$-simplex

$$\Delta_m = \left\{ (p_0, \ldots, p_m) \mid p_i \in (0,1), \sum_{i=0}^{m} p_i = 1 \right\}.$$

**Definition.**
An algebraic variety $X$ is **toric** if it contains a dense algebraic torus whose action on itself extends to $X$.

**Normal toric varieties. . .**
. . . of dimension $n$ are encoded by complete fans in $\mathbb{R}^n$. 
Example: a coin model

\( \mathbf{X} = \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \)  
\( \mathbf{X} = \mathbf{X}_\Sigma \)  
\( \mathbf{X}_{>0} \cong (0, 1)^3 \)

\( \)homogeneous coordinates \((x_0 : x_1), (s_0 : s_1), (t_0 : t_1)\)
\( \)\( \Sigma \) the inner normal fan of \([0, 1]^3 \)
\( \)the positive part of \( \mathbf{X} \)

**Model:**  
image of \( \mathbf{X}_{>0} \rightarrow \Delta_m \), \((x, s, t) \mapsto (p_\ell(x, s, t))_{\ell=0, \ldots, m}, \)
\( x = x_0, x_1 = 1 - x, s = s_0, s_1 = 1 - s, t = t_0, t_1 = 1 - t \)

\[
p_\ell = \binom{m}{\ell} x^\ell s^{m-\ell} + \binom{m}{\ell} (1-x)^\ell t^{m-\ell}, \quad \ell = 0, 1, \ldots, m.
\]

*probability for observing \( \ell \) times head*

**Marginal likelihood integral**

For uniform prior on \((0, 1)^3 \), data \( u = (u_0, \ldots, u_m) \), the **marginal likelihood integral** is

\[
\mathcal{I}_u = \int_{\mathbf{X}_{>0}} p_{u_0}^u \cdots p_{u_m}^u \mathcal{O}_{\mathbf{X}}^{\text{unif}}.
\]

\( u_+ = u_0 + \cdots + u_m \) many repetitions
Normal toric varieties

Example: complex projective plane

$\Sigma$ the inner normal fan of $\Delta_2$, $V = (v_1|v_2|v_3) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$

\[ X_\Sigma = \mathbb{P}^2 = (\mathbb{C}^3)^*/\mathbb{C}^* = (\mathbb{A}^3_\mathbb{C} \setminus \mathcal{V}(x_0, x_1, x_2))/\mathbb{G}^1_m \]

Homogeneous coordinates: $(x_0 : x_1 : x_2)$  
Cox coordinates

Three affine charts: $\{ x_i \neq 0 \}$  
one for each maximal cone

In general

$\Sigma$ a complete fan in $\mathbb{R}^n$  
e.g. the inner normal fan of a polytope $P$

- $V = (v_1|\cdots|v_k)$  
columns: primitive ray generators of the $\rho_i \in \Sigma(1)$
- $\text{Cl}(X) = \mathbb{Z}^k / \text{im}(V^\top)$  
divisor class group of $X$
- $G = \text{Hom}(\text{Cl}(X), \mathbb{C}^*)$  
the characters of $\text{Cl}(X)$
- $S = \mathbb{C}[x_1, \ldots, x_k] = \bigoplus_{\gamma \in \text{Cl}(X)} S_\gamma$  
Cox ring
- $B = \langle \prod_{\rho \notin \sigma} x_\rho \mid \sigma \in \Sigma(n) \rangle \subset S$  
the irrelevant ideal
- $X_\Sigma = (\mathbb{C}^k \setminus \mathcal{V}(B))/G$  
the toric variety of $\Sigma$

Positive toric varieties

Setup

\[ \Sigma \text{ the inner normal fan of a polytope } P \]
\[ X = X_\Sigma \text{ the toric variety of } \Sigma \]
\[ P^\circ \text{ the interior of } P \]

Positive part of \( X_\Sigma = (\mathbb{C}^k \setminus \mathcal{V}(B)) / G \)

\[ \blacklozenge \pi: \mathbb{C}^k \setminus \mathcal{V}(B) \rightarrow (\mathbb{C}^k \setminus \mathcal{V}(B)) / G \text{ the projection} \]
\[ \blacklozenge \pi(\mathbb{R}^k_{>0}) =: X_{>0} \text{ the positive part of } X_\Sigma \quad X_{\geq 0} \text{ its Euclidean closure} \]

Algebraic moment map

One identifies \( X_{>0} \) and \( P^\circ \) via the moment map

\[ X_{>0} \xrightarrow{\phi} \mathbb{R}^n_{>0} \xrightarrow{\varphi} P^\circ, \]

with \( \varphi \) the affine moment map

\[ \varphi(t) = \sum_{a \in \mathcal{V}(P)} \frac{c_a t^a}{q(t)} \cdot a, \quad q = \sum_{a \in \mathcal{V}(P)} c_a t^a \in \mathbb{R}_{>0}[t_{1}^{\pm 1}, \ldots, t_{n}^{\pm 1}]. \]
The **canonical form** of \((X, X_{\geq 0})\) is the meromorphic differential \(n\)-form

\[
\Omega_X = \sum_{l \subseteq \Sigma(1), |l| = n} \det(V_l) \bigwedge_{\rho \in l} \frac{dx_\rho}{x_\rho}
\]

on \(X\). The pair \((X, X_{\geq 0})\) is a *positive geometry*.

**Proposition**

The pullback of \(dy_1 \wedge \cdots \wedge dy_n\) on \(P^\circ\) under the moment map \(X_{\geq 0} \to P^\circ\) is a positive rational function \(r\) times \(\Omega_X\). We obtain \(r(x)\) from \(|\det|\) of the *toric Hessian* of \(\log(q(t))\)

\[
H(t) = (\theta_i \theta_j \cdot \log(q(t)))_{i,j} \quad \theta_i = t_i \partial_{t_i}
\]

by replacing \(t_1, \ldots, t_n\) with Laurent monomials in \(x_1, \ldots, x_k\) given by the rows of \(V\).

**Observation:** Scaled by a rational function \(\frac{f}{g}\), \(\Omega_X\) gives a probability measure on \(X_{\geq 0}\).

**Integrals of interest:** \(\mathcal{I}_{f,g} = \int_{X_{\geq 0}} \frac{f}{g} \Omega_X\) \(f, g \in S\) homogeneous of the same degree

Toric sector decomposition

Definition

The **tropical approximation** of \( f \in \mathbb{C}[x_1, \ldots, x_k] \) is the piecewise monomial function

\[
    f^{\text{tr}} : \mathbb{R}^k_{>0} \longrightarrow \mathbb{R}_{>0}, \quad x \mapsto \max_{\ell \in \text{supp}(f)} x^\ell.
\]

Proposition

Let \( \mathcal{F} \) be a simplicial refinement of the normal fan of \( \mathcal{N}(f) + \mathcal{N}(g) \). Then

\[
    \mathcal{I}_{f,g} = \int_{X_{>0}} \frac{f}{g} \Omega_X = \sum_{\sigma \in \mathcal{F}(n)} \int_{\text{Exp}(\sigma)} \frac{f^{\text{tr}}}{g^{\text{tr}}} = \sum_{\sigma \in \mathcal{F}(n)} \mathcal{I}_\sigma \quad \text{sector integrals}
\]

\( \diamond \ \text{Exp} : \mathbb{R}^k/K \rightarrow X_{>0}, \ [y_1, \ldots, y_k] \mapsto \pi(e^y) \)

\( \diamond \ \text{parameterization} \ x^\sigma : [0,1]^n \rightarrow \text{Exp}(\sigma) \)

Tropical detour

Also the tropical integral \( \mathcal{I}_{f,g}^{\text{tr}} = \int_{X_{>0}} f^{\text{tr}}/g^{\text{tr}} \Omega_X \) decomposes as \( \mathcal{I}_{\text{tr}} = \sum_{\sigma \in \mathcal{F}(n)} \mathcal{I}_{\sigma}^{\text{tr}} \).

Each tropical sector integral \( \mathcal{I}_{\sigma}^{\text{tr}} \) is an integral over a monomial encoded by data of \( \mathcal{F}! \)

\[
    \mathcal{I}_{\sigma}^{\text{tr}} = \int_{\text{Exp}(\sigma)} x^{-(\nu_g - \nu_f)} \Omega_X
\]
Toric data of $f, g$

**Theorem**
Suppose that the Newton polytope of $g$ is $n$-dimensional and contains that of the numerators $f$ in its relative interior. Then the integral $\int_{X > 0} f/g \, \Omega_X$ converges.

**Proposition**
Let $\mathcal{F}$ a simplicial refinement of $\mathcal{N}(f) + \mathcal{N}(g)$. Let $\sigma$ be a cone of $\mathcal{F}$, $\nu_f$ and $\nu_g$ corresponding faces of $\mathcal{N}(f)$ and $\mathcal{N}(g)$. Then:

$$\frac{f^{\text{tr}(x)}}{g^{\text{tr}(x)}} = x^{-(\nu_g - \nu_f)}$$

for all $x \in \mathbb{R}^k$ such that $\pi(x) \in \text{Exp}(\sigma)$.

Then

$$I^{\text{tr}} = \sum_{\sigma \in \mathcal{F}(n)} I^{\text{tr}}_{\sigma}$$

where

$$I^{\text{tr}}_\sigma = \int_{\text{Exp}(\sigma)} \frac{f^{\text{tr}}}{g^{\text{tr}}} \Omega_X = \int_{\text{Exp}(\sigma)} x^{-(\nu_g - \nu_f)} \Omega_X.$$

Write $\text{im}(V^\top)$ as $\ker(W)$, $W = (w_1| \cdots |w_n)$. The tropical sector integral is equal to

$$I^{\text{tr}}_\sigma = \frac{\det(VW)}{\prod_{\ell=1}^n w_\ell \cdot (\nu_g - \nu_f)}.$$
Sampling from \( (X_{>0}, d_{f,g}^{(tr)}) \)

\[
\mu_{f,g} = \frac{1}{\mathcal{I}_{f,g}} \cdot \frac{f}{g} \Omega_X \quad \text{and} \quad \mu_{f,g}^{tr} = \frac{1}{\mathcal{I}_{f,g}^{tr}} \cdot \frac{f^{tr}}{g^{tr}} \Omega_X \quad \text{are probability measures on } X_{>0}!
\]

Sampling from the tropical density

**Input:** \( \mathcal{F}, \mathcal{I}_{f,g}^{tr}, \text{ and } \mathcal{I}^{tr}. \)

**Step 1.** Draw an \( n \)-dimensional cone \( \sigma \) from \( \mathcal{F}(n) \) with probability \( \mathcal{I}_{\sigma}^{tr} / \mathcal{I}^{tr} \).

**Step 2.** Draw a sample \( q \) from the unit hypercube \([0,1]^n\) using the uniform distribution.

**Step 3.** Compute \( x^{\sigma}(q) \in X_{>0} \).

**Output:** The element \( x^{\sigma}(q) \in X_{>0}, \) a sample from \( (X_{>0}, d_{f,g}^{tr}) \).

*Sampling from \( d_{f,g} \) via rejection sampling!*

**Proposition**

Let \( x^{(1)}, \ldots, x^{(N)} \) be tropical samples from \( X_{>0} \). Then

\[
h(x) = \frac{f(x) \cdot g^{tr}(x)}{g(x) \cdot f^{tr}(x)}
\]

\[
\mathcal{I}_{f,g} \approx \mathcal{I}_N = \frac{\mathcal{I}_{f,g}^{tr}}{N} \cdot \sum_{i=1}^{N} h\left(x^{(i)}\right).
\]
Bayesian inference

Toric polytope models \( c = (c_0, \ldots, c_m), \ c_i \in \mathbb{R}_{>0} \)

- \( Z = c_0 x^{a_0} + c_1 x^{a_1} + \cdots + c_m x^{a_m} \in S \) homogeneous of degree \( \gamma \in \text{Cl}(X) \)
  - \( a_i \) lattice points of \( P \)
- \( p_i = c_i x^{a_i} / Z, \ i = 0, \ldots, m, \) are positive on \( X_{>0}, \ \sum_{i=0}^{m} p_i = 1 \)
  - statistical model: image of resulting map \( X_{>0} \rightarrow \Delta_m \)

Bayes’ factor for toric pentagon model

Prior: distribution \( \mu_{f,g} \) arising from uniform distribution on \( P^\circ \)
Data: \( u = (u_0, \ldots, u_5) = (1, 2, 4, 8, 16, 32) \quad u_+ = \sum u_i = 63 \)

Competing models: toric models \( \mathcal{M}_c \) for

\( c^{(1)} = (2, 3, 5, 7, 11, 13) \quad \text{and} \quad c^{(2)} = (32, 16, 8, 4, 2, 1). \)

Marginal likelihood integrals:

\[
\mathcal{I}_{u}^{(i)} = \int_{X_{>0}} \left( \frac{L_u^{(i)}(x)}{x^u} \right) \mu_{f,g}, \quad i = 1, 2.
\]

Bayes’ factor: \( K = \mathcal{I}_{u}^{(1)}/\mathcal{I}_{u}^{(2)} \approx 21.06. \)

\( \mathcal{M}_{c^{(1)}} \) is a better fit for the data than \( \mathcal{M}_{c^{(2)}}! \)
Sampling from \((X_{>0}, d_{f,g})\)

Setup

- \(d_1\) and \(d_2\) two densities on the same space with the same differential form
e.g. on \((X_{>0}, \Omega_X)\)
- suppose it is hard to sample from \(d_1\), but easy to sample from \(d_2\)
- suppose there exists \(C \geq 1\) such that \(d_1(x)/d_2(x) \leq C\) for all \(x\)

Rejection sampling

Step 1. Draw a sample \(x \in X\) using \(d_2\), and \(\xi \in [0, C]\) with the uniform distribution.
Step 2. If \(\xi < d_1(x)/d_2(x)\), accept \(x\) as a sample. Otherwise, reject \(x\).

Output: A sample from \(d_2(x) \cdot d_1(x)/d_2(x)\), i.e., \(d_1(x)\).

Proposition

Suppose that \(f = \sum_{\ell \in \text{supp}(f)} f_\ell x^\ell\) has positive coefficients. Set \(C_1 = \min_{\ell \in \text{supp}(f)} f_\ell\) and \(C_2 = \sum_{\ell \in \text{supp}(f)} f_\ell\). Then

\[
0 < C_1 \leq \frac{f(x)}{f_{tr}(x)} \leq C_2 < \infty \quad \text{for all } x \in X_{>0}.
\]

Sampling from \(d_{f,g}\) via rejection sampling with \(d_{f,g}^{tr}\)!
In a nutshell

1. Statistical models parameterized by toric varieties occur naturally.
2. Positive toric varieties are probability spaces. \( \text{positive geometries} \)
3. Bayesian inference via tropical methods. \( \int_{X > 0} L_u \Omega^\text{prior}_X, \int_{X > 0} f/g \Omega_X \)

Supplementary material

- code in Julia available at: https://mathrepo.mis.mpg.de/BayesianIntegrals
- painting inspired by the pentagon model: https://alsattelberger.de/painting/

Thank you for your attention!
Error estimates

Let \( h(x) = \frac{f(x) \cdot g^{\text{tr}}(x)}{g(x) \cdot f^{\text{tr}}(x)} \). Then

\[
M_1 \leq h(x) \leq M_2 \quad \text{for all } x \in X_{>0},
\]

where

\[
M_1 = \frac{\min_{\ell \in \text{supp}(f)} f_\ell}{\sum_{\ell \in \text{supp}(g)} g_\ell} \quad \text{and} \quad M_2 = \frac{\sum_{\ell \in \text{supp}(f)} f_\ell}{\min_{\ell \in \text{supp}(g)} g_\ell}.
\]

Proposition

Let \( x^{(1)}, \ldots, x^{(N)} \) be tropical samples from \( X_{>0} \). Then

\[
\mathcal{I}_{f,g} \approx \mathcal{I}_N = \frac{\mathcal{I}_{f,g}^{\text{tr}}}{N} \cdot \sum_{i=1}^{N} h \left( x^{(i)} \right).
\]

Proposition

The standard deviation of the approximation above satisfies

\[
\sqrt{\mathbb{E} \left[ (\mathcal{I} - \mathcal{I}_N)^2 \right]} \leq \mathcal{I}^{\text{tr}} \cdot \sqrt{\frac{M_2^2 - M_1^2}{N}}.
\]
The Wachspress model

$P \subset \mathbb{R}^n$ a polytope, $\Sigma$ its inner normal fan, $V = (v_1| \cdots |v_k)$

Inequality representation of $P$

$$P = \{ y \in \mathbb{R}^n | \langle v_i, y \rangle + \alpha_i \geq 0, \ i = 1, 2, \ldots, k \}$$

with $\alpha_1, \ldots, \alpha_k \in \mathbb{Z}_{>0}$. The vertices $q_I$ of $P$ are indexed by cones $I \in \Sigma(n)$: the vertex $q_I \in \mathbb{Z}^n$ is the unique solution of $\langle v_i, y \rangle = -\alpha_i$ for $i \in I$.

Definitions

The adjoint of $P$ is the polynomial in variables $y_1, \ldots, y_n$

$$A = \sum_{I \in \Sigma(n)} |\det (\tilde{V}_I)| \cdot \prod_{i \notin I} (1 + \frac{1}{\alpha_i} \langle v_i, y \rangle) .$$

The Wachspress model of $P$ is the image of $P \rightarrow \Delta_m$, $y \mapsto (p_I(y))_{I \in \Sigma(n)}$ with

$$p_I(y) = \frac{|\det (\tilde{V}_I)|}{A(y)} \cdot \prod_{I \in \Sigma(n)} \left(1 + \frac{1}{\alpha_i} \langle v_i, y \rangle \right).$$