Bayesian Lipschitz Constant Estimation and Quadrature

Jan-Peter Calliess

Department of Engineering
University of Cambridge

jpc73@cam.ac.uk

December 11, 2015

Overview

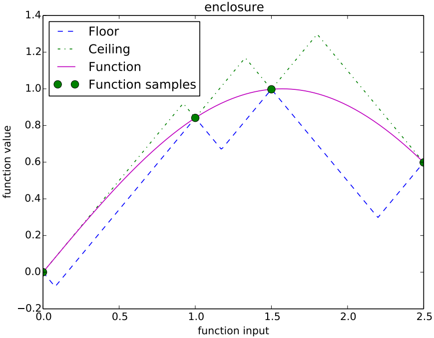
1. Lipschitz quadrature
2. Bayesian Lipschitz constant estimation method
3. Combination: Bayesian Lipschitz quadrature
4. Extensions and future work

Lipschitz quadrature

Recap: Lipschitz continuity

f Lipschitz with constant $L \Leftrightarrow \forall x, x' : |f(x) - f(x')| \leq L\,\mathfrak{d}(x, x')$.

- Given: sample $\mathcal{D}$ of integrand f.
- f has Lipschitz constant $L \Rightarrow$ can compute integrals of bounds $\mathfrak{u}(\cdot; L)$ ("ceiling") and $\mathfrak{l}(\cdot; L)$ ("floor") (e.g. [Baran et. al . '08]).



- $\int_I \mathfrak{l}(x; L)\, \mathrm{d}x \leq \int_I f(x)\, \mathrm{d}x \leq \int_I \mathfrak{u}(x; L)\, \mathrm{d}x$.
- $\mathfrak{l}, \mathfrak{u} \overset{unif.}{\rightarrow} f$ as the data becomes dense [Calliess '14].

Lipschitz quadrature

Recap: Lipschitz continuity

f Lipschitz with constant $L \Leftrightarrow \forall x, x' : |f(x) - f(x')| \leq L \, \mathfrak{d}(x, x')$.

Practical problem:

- Lipschitz constant L may be unknown a priori.

Task:

- Infer subjective belief over (the smallest) L from a sample.
- Fold in uncertainty into integral estimates.
- Obtain Bayesian uncertainty bounds around integral estimates.

Bayesian inference over the best Lipschitz constant

Recap: Lipschitz continuity

f Lipschitz with constant $L \Leftrightarrow \forall x, x' : |f(x) - f(x')| \leq L\, \mathfrak{d}(x, x')$.

Best Lipschitz constant

- $L^* := \sup_{x,x', \mathfrak{d}(x,x')>0} R(x, x') \leq L$, where $R(x, x') = \frac{|f(x)-f(x')|}{\mathfrak{d}(x,x')}$.
- $L^* =$ "best" (i.e. smallest) Lipschitz constant.

Bayesian inference:

Calculate posterior: $\pi(L^* = \ell|\mathcal{D}) = \frac{\pi(\mathcal{D}|L^*=\ell)\pi_0(L^*=\ell)}{\int_0^\infty \pi(\mathcal{D}|L^*=\ell)\pi_0(L^*=\ell)\, \mathrm{d}\ell}$.

Bayesian inference over the best Lipschitz constant

Best Lipschitz constant

$L^* := \sup_{x,x',\, \mathfrak{d}(x,x')>0} R(x,x')$, where $R(x,x') = \frac{|f(x)-f(x')|}{\mathfrak{d}(x,x')}$.

L^* as a r.v. and its likelihood

Inputs $x_i \sim$ Unif. $\rightarrow$ data $\mathcal{D} = \{\big(x_i, f(x_i)\big)|i = 1, ..., n\}$
$\rightarrow S := \{R(x_i, x_j)|i < j,\, \mathfrak{d}(x_i, x_j) > 0\}$ set of r.v.
with $R(x_i, x_j) \sim \text{Unif}(0, L^*) \rightarrow \pi(S|L^*)$.

Bayesian inference:

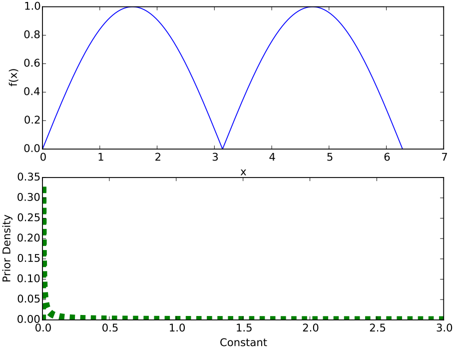
$\pi(L^* = \ell|\mathcal{D}) \overset{!}{=} \pi(L^*|S) = \frac{\pi(S|L^*)\pi_0(L^*)}{\int_0^\infty \pi(S|L^*)\pi_0(L^*)\, \mathrm{d}L^*}$.

Bayesian inference over the Lipschitz constant

- We place a Pareto prior over the best Lipschitz constant...

Bayesian belief:

- Prior: $Pa(b, K)$.
- Uninformative prior: limit case $\mu, b \to 0$.
- Pareto prior and uniform likelihood are conjugate $\to$ closed-form inference.

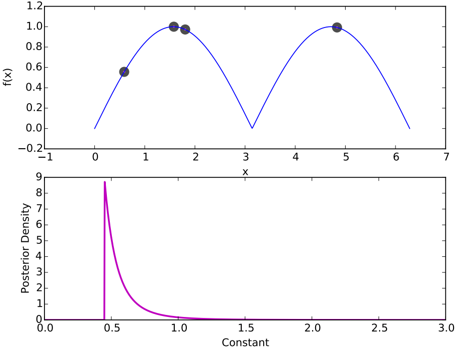


Bayesian inference over the Lipschitz constant

- We place a Pareto prior over the best Lipschitz constant...
- ... we observe a *(small)* sample $\mathcal{D}$ and update our subjective belief...

Bayesian belief:

- Posterior: $\pi(L^*|S) = \frac{\pi(S|L^*)\pi_0(L^*)}{\int_0^\infty \pi(S|L^*)\pi_0(L^*)\, \mathrm{d}L^*}$.
- Pareto prior and uniform likelihood are conjugate $\rightarrow$ closed-form inference.

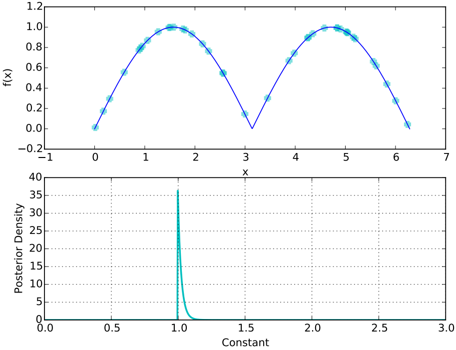


Bayesian inference over the Lipschitz constant

- We place a Pareto prior over the best Lipschitz constant...
- ... we observe a *(larger)* sample $\mathcal{D}$ and update our subjective belief...

Bayesian belief:

- Posterior: $\pi(L^*|S) = \frac{\pi(S|L^*)\pi_0(L^*)}{\int_0^\infty \pi(S|L^*)\pi_0(L^*)\,\mathrm{d}L^*}$.
- Pareto prior and uniform likelihood are conjugate $\rightarrow$ closed-form inference.

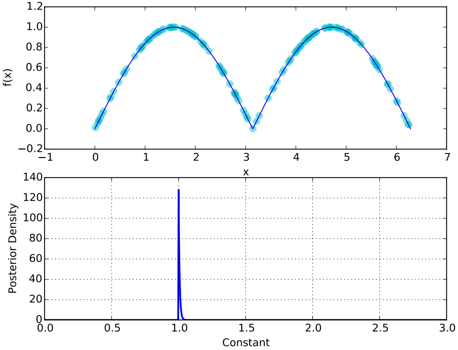


Bayesian inference over the Lipschitz constant

- We place a Pareto prior over the best Lipschitz constant...
- ... we observe a *(-n even larger)* sample $\mathcal{D}$ and update our subjective belief...

Bayesian belief:

- Posterior: $\pi(L^*|S) = \frac{\pi(S|L^*)\pi_0(L^*)}{\int_0^\infty \pi(S|L^*)\pi_0(L^*)\,\mathrm{d}L^*}$.
- Pareto prior and uniform likelihood are conjugate $\rightarrow$ closed-form inference.



Integral estimate with Bayesian confidence bounds

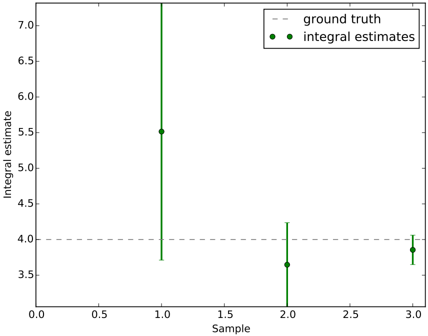
- For $\theta \in (0,1)$, **find**:
  $L_\theta = \min\{L \geq 0 | \int_0^L \pi(L^* = \ell|\mathcal{D}) \, \mathrm{d}\ell \geq 1-\theta\}$.
- With probability at least $1-\theta$ the true smallest Lipschitz constant is less than or equal to L_θ.
- Using this L_θ as the Lipschitz constant parameter in Lipschitz quadrature implies:

Bayesian confidence bounds on the integral:

- $\Pr\left(\int_I f(x) \, \mathrm{d}x \in \left[S_l(L_\theta), S_u(L_\theta)\right]\right) \geq 1-\theta$
- where $S_l(L) = \int_I \mathfrak{l}_n(x;L) \, \mathrm{d}x$ and $S_u(L) = \int_I \mathfrak{u}_n(x;L) \, \mathrm{d}x$ are the integrals of the upper and lower bound functions, respectively.

Integral estimate with Bayesian confidence bounds

- Integral bounds for our examples with confidence parameter $\theta = 0.001$.



Extensions and Future Work

- $\pi(L^* = \ell|\mathcal{D}) \stackrel{!?}{=} \pi(L^* = \ell|S)$.
- Cubature.
- Indefinite integrals.
- Hoelder continuous functions.
- Applications to Bayesian risk estimates in sparse data scenarios.
- Optimisation and control.
- Local Lipschitz continuity and constant estimates.
- Suggestions?