Sparse parametric estimation of Poisson processes

Michael G. Moore and Mark A. Davenport
Georgia Institute of Technology
{mmoore90,mdav}@gatech.edu

I. Introduction

In this work we provide a recovery guarantee for estimating the parameters of an inhomogeneous Poisson arrival processes where event arrival rates are dictated by a weighted combination of known functions. Specifically, we consider the problem of estimating the parameters $\bar{x} \in \mathbb{R}^N$ of a Poisson process with rate

$$R_{\bar{x}}(t) = g(t) + \sum_{n=1}^{N} \bar{x}_n \gamma_n(t) \tag{1}$$

supported for $t \in \mathbb{T}$ with known real-valued functions $g(t)$ and $\gamma_n(t)$. We base our estimate on a set of observed event times τ drawn from this process, i.e., for any $T \subseteq \mathbb{T}$ the event times τ satisfy

$$|\tau \cap T| \sim \text{Poisson}\left(\int_T R_{\bar{x}}(t)dt\right). \tag{2}$$

Previous results have achieved reliable Poisson estimation mostly by adapting existing estimators intended for independent noise (e.g., [1]) and studied this problem from the perspective of minimax risk bounds that hold for arbitrary estimators [2]. Only recently have recovery guarantees been established for the natural maximum-likelihood estimator, which often outperforms other estimators in practice [3]. However, these results apply only to Poisson counting processes, in which event arrivals are histogrammed into bins (or, equivalently in our framework, the rate function $R_{\bar{x}}(t)$ is piecewise-constant).[1]

The guarantee we provide here generalizes parameter estimation for Poisson counting processes to the more general setting of continuous-time Poisson arrival processes (i.e., with infinite arrival resolution). We provide a general guarantee but highlight that improved results are possible when the solution is known to be sparse, as is already known in the Poisson counting case. Additionally, we note that our result requires fewer assumptions than previous results for Poisson estimation. In particular, we do not require the common constraint that the rate be a nonnegative combination of nonnegative functions and our guarantee holds without oracle knowledge regarding the true parameters.

II. Recovery guarantee

Before stating our main result, we briefly fix some notation. The negative log-likelihood of a set of observations τ under a candidate parameter set x is

$$\mathcal{L}(\tau|x) = \int_{\mathbb{T}} R_x(t)dt - \sum_{m=1}^{|\tau|} \log R_x(\tau_m). \tag{3}$$

Our guarantee will depend on bounds $R_{\min}$ and $R_{\max}$ such that $R_{\min} \leq R_{\bar{x}}(t) \leq R_{\max}$. Letting Σ_k^N be the set of all subsets of $\{1 \dots N\}$ of cardinality at most k, we make the definitions

$$\Gamma_{ij} = \int_{\mathbb{T}} \gamma_i(t)\gamma_j(t)dt \qquad \gamma_k = \max_{s \in \Sigma_k^N} \sup_{t \in \mathbb{T}} \sqrt{\sum_{n \in s} \gamma_n^2(t)}.$$

We also say that $\{\gamma_n(t)\}$ satisfies the restricted isometry property with constant δ_k if

$$(1-\delta_k)\|x\|_2^2 \leq x^T \Gamma x \leq (1+\delta_k)\|x\|_2^2 \quad \forall \|x\|_0 \leq k. \tag{4}$$

Using these definitions, we can state our main result as follows.

Theorem 1: Let c_α be a value that depends only on a bound $\alpha \geq \frac{\zeta \gamma_k^2}{R_{\min} k (1+\delta_k)}$. If we require that $R_{\min} \geq \frac{c_\alpha^2 k \zeta \gamma_k^2 (1+\delta_k)}{(1-\delta_k)^2}$ then any vector $\widehat{x}$ satisfying $\|\widehat{x} - \bar{x}\|_0 \leq k$ and $\mathcal{L}(\tau|\widehat{x}) \leq \mathcal{L}(\tau|\bar{x})$ will also satisfy

$$\|\widehat{x} - \bar{x}\|_2 \leq c_\alpha \frac{\sqrt{k\zeta(1+\delta_k)}}{1-\delta_k} \frac{R_{\max}}{\sqrt{R_{\min}}} \tag{5}$$

with probability at least $1 - (2k+3)\exp(-\zeta)$.

Note that this theorem also yields a corollary for recovery based on Poisson bin counts, where observations instead take the form $y \sim \text{Poisson}(g + A\bar{x})$ for a known vector g and RIP matrix A. The corollary is trivially realized by restricting the functions $g(t)$ and $\gamma_n(t)$ to be piecewise-constant.

III. Discussion

The constrained maximum likelihood estimator

$$\widehat{x} = \arg\min_{x \in \mathcal{X}} \mathcal{L}(\tau|x) \tag{6}$$

will satisfy the theorem for any set $\mathcal{X}$ such that $\bar{x} \in \mathcal{X}$. If $\mathcal{X}$ is convex then the estimator (6) is also convex and is thus amenable to convex optimization techniques (e.g., [4, 5]).

Taking advantage of sparsity ($k < N$) with tractable programs introduces some difficulties. One can choose the nonconvex set $\mathcal{X} = \{x : \|x\|_0 \leq \|\bar{x}\|_0\}$ and satisfy the theorem with $k \leq 2\|\bar{x}\|_0$, but the optimization problem is combinatorial. Another option is $\mathcal{X} = \{x : \|x\|_1 \leq \|\bar{x}\|_1\}$, for which a value $k \leq \|\bar{x}\|_0 + \|\widehat{x}\|_0$ can be used. While this set is convex and encourages sparse solutions, directly applying the results of Theorem 1 requires a guarantee controlling the cardinality k of $\widehat{x}$. We leave such a guarantee to future work.

Finally, we also note that it is possible to increase $R_{\min}$ arbitrarily by increasing $g(t)$. This can be accomplished by artificially adding events from a homogeneous Poisson process to the observations τ. If we add events with rate $R_{\max}$, we can change the $R_{\max}/\sqrt{R_{\min}}$ scaling to $\sqrt{R_{\max}}$, removing any dependence on the dynamic range. While this is an interesting theoretical improvement, we suspect that this additional noise is likely to degrade performance in practice.

Acknowledgments

This work was supported by grants NRL N00173-14-2-C001, AFOSR FA9550-14-1-0342, NSF CCF-1409406, CCF-1350616, and CMMI-1537261.

[1] It might seem possible to apply existing results merely by driving the "bin widths" to zero, but for technical reasons this leads to degenerate results.

REFERENCES

[1] X. Jiang, P. Reynaud-Bouret, V. Rivoirard, L. Sansonnet, and R. Willett, "A data-dependent weighted LASSO under Poisson noise," *arXiv:1509.08892*, Sept. 2015.

[2] X. Jiang, G. Raskutti, and R. Willett, "Minimax optimal rates for Poisson inverse problems with physical constraints," *IEEE Trans. Inform. Theory*, vol. 61, no. 8, pp. 4458–4474, Aug. 2015.

[3] M. Rohban, V. Saligrama, and D. Vaziri, "Minimax optimal sparse signal recovery with Poisson statistics," *IEEE Trans. Signal Processing*, vol. 64, no. 13, pp. 3495–3508, July 2016.

[4] Z. Harmany, R. Marcia, and R. Willett, "This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms – Theory and Practice," *IEEE Trans. Image Processing*, vol. 21, no. 3, pp. 1084–1096, 2012.

[5] Q. Tran-Dinh, A. Kyrillidis, and V. Cevher, "Composite self-concordant minimization," *J. Machine Learning Research*, vol. 16, pp. 371–416, 2015.