Functional Representation Lemma: Algorithms and Applications

Author(s):
Shkel, Yanina

Publication date:
2024-03-06

Permanent link:
https://doi.org/10.3929/ethz-b-000664570

Functional Representation Lemma: Algorithms and Applications

Yanina Shkel
École Polytechnique Fédérale de Lausanne (EPFL)
email: yanina.shkel@epfl.ch

***Abstract*—Functional Representation Lemma (FRL) is an information-theoretic technique that fixes a correlated 'reference' information source, and extracts a 'residual' information about the original source. Recently, there has been a lot of interest in FRL since variants of this technique appear across different problems in information theory, and data science more broadly.**

In this tutorial talk we overview the FRL problem. We highlight some of its applications: these include the problems of privacy and causal inference, as well as proof techniques for single-shot information-theoretic bounds. Finally, we review known algorithms for constructing functional representations. We particularly focus on the greedy algorithms previously proposed in literature.

I. Extended Abstract

A. Overview and Applications

We begin with the Simple Functional Representation Lemma (FRL) which can be found in [1] and was independently derived in [2]–[4], among others. Given two jointly distributed discrete random variable (X, Y), the lemma states that there exists a random variable Z such that

$$I(Y;Z) = 0, \tag{1}$$
$$H(X|Z,Y) = 0, \tag{2}$$
$$\text{and } |\mathcal{Z}| \leq |\mathcal{Y}|(|\mathcal{X}| - 1) + 1. \tag{3}$$

That is, Y and Z are independent, X is a deterministic function of Y and Z, and the support of Z is bounded. This result could be shown with a construction that we here call the *Simple FRL algorithm*. See, for example, [1, Appendix B] and [4, Lemma 1] for a detailed exposition.

The simple FRL shows that a random variable Z that satisfies (1) and (2) exists. However, there are many more interesting questions that arise about properties of this random variable. One line of work focuses on minimizing $H(X|Z)$ (or maximizing $I(X;Z)$). The best known result, known as the Strong Functional Representation Lemma (SFRL), states that it is possible to construct Z, such that

$$H(X|Z) \leq I(X;Y) + \log\left(I(X;Y) + 1\right) + O(1), \tag{4}$$

where $I(X;Y)$ is a trivial lower bound on $H(X|Z)$ [5], [6]. An extensions of SFRL, known as the Poisson Matching Lemma, has been proposed in [7]. These results find extensive applications in derivations of single-shot coding bounds [5]–[8], as well as for problems in information-theoretic privacy [4], [9].

Another line of work on minimizing the entropy $H(Z)$ finds applications in the problem of private compression [4], [15], causal inference [16]–[20], as well as a number of other problems in statistics [21]. Let $Q = \bigwedge_{y\in\mathcal{Y}} P_{X|Y=y}$ be the lower bound with respect to majorization [21] of the set of distributions $\{P_{X|Y=y}\}_{y\in\mathcal{Y}}$. It can be shown that

$$H\left(Q\right) \leq H(Z) \leq H\left(Q\right) + 2 - 2^{2-|\mathcal{Y}|}. \tag{5}$$

The lower bound in (5) was shown in [21]. It was also shown in [21] that the upper bound for $|\mathcal{Y}| = 2$ holds via a greedy algorithms that we refer to as the *majorization-based algoirhtm*. The general upperbound in (5) was shown in [22] using the technique of *geometric splitting*.

An improvement on (5) has been recently shown using the information spectrum of Z. Specifically [23], [24] show that

$$\mathbb{P}[\imath_Z(Z) > t] \geq \sup_{y\in\mathcal{Y}} \mathbb{P}[\imath_{X|Y}(X|Y) > t|Y = y] \tag{6}$$

where $\imath_Z(z) = \log\frac{1}{P_Z(z)}$ and $\imath_{X|Y}(x|y) = \log\frac{1}{P_{X|Y}(x|y)}$. Moreover, [23], [24] show that there exists a distribution Q^* such that

$$H(Q) \leq H(Q^*) \leq H(Z). \tag{7}$$

This Q^* could be found with a simple greedy algorithm from the information spectrum envelope on the right-hand-side of (6). Finally, [23], [25] show that an algorithm that we call the *natural greedy algorithm* is within $\frac{\log_2(e)}{e} \approx 0.53$ bits of the minimal achievable entropy, while, in general, the problem is known to be NP-hard.

B. On Greedy Algorithms

In this talk, we particularly focus on greedy algorithm for the problem of constructing Z. This includes the the majorization-based algorithm which attempts to best approximate the greatest lowebound Q in (5). The natural greedy algorithm, on the other hand, puts as much probability mass as possible into the likelier realizations of Z. Greedy algorithms do not just play a role with constructing the random variable Z. The also show up in the evaluations of lower bounds in (5) and (7).

REFERENCES

[1] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[2] B. Hajek and M. Pursley, "Evaluation of an achievable rate region for the broadcast channel," *IEEE Transactions on Information Theory*, vol. 25, no. 1, pp. 36–46, January 1979.

[3] F. Willems and E. van der Meulen, "The discrete memoryless multiple-access channel with cribbing encoders," *IEEE Transactions on Information Theory*, vol. 31, no. 3, pp. 313–327, 1985.

[4] Y. Y. Shkel, R. S. Blum, and H. V. Poor, "Secrecy by design with applications to privacy and compression," *IEEE Transactions on Information Theory*, vol. 67, no. 2, pp. 824–843, 2021.

[5] C. T. Li and A. E. Gamal, "Strong functional representation lemma and applications to coding theorems," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 589–593.

[6] P. Harsha, R. Jain, D. McAllester, and J. Radhakrishnan, "The communication complexity of correlation," *IEEE Trans. Inf. Theor.*, vol. 56, no. 1, pp. 438–449, jan 2010. [Online]. Available: https://doi.org/10.1109/TIT.2009.2034824

[7] C. T. Li and V. Anantharam, "A unified framework for one-shot achievability via the poisson matching lemma," *IEEE Transactions on Information Theory*, vol. 67, no. 5, pp. 2624–2651, 2021.

[8] L. Theis and A. B. Wagner, "A coding theorem for the rate-distortion-perception function," 2021.

[9] A. Zamani, T. J. Oechtering, and M. Skoglund, "On the privacy-utility trade-off with and without direct access to the private data," *IEEE Transactions on Information Theory*, pp. 1–1, 2023.

[10] A. Makhdoumi, S. Salamatian, N. Fawaz, and M. Médard, "From the information bottleneck to the privacy funnel," in *2014 IEEE Information Theory Workshop (ITW 2014)*, 2014, pp. 501–505.

[11] F. d. P. Calmon, A. Makhdoumi, M. Médard, M. Varia, M. Christiansen, and K. R. Duffy, "Principal inertia components and applications," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 5011–5038, Aug 2017.

[12] S. Asoodeh, M. Diaz, F. Alajaji, and T. Linder, "Information extraction under privacy constraints," *Information*, vol. 7, no. 1, 2016. [Online]. Available: https://www.mdpi.com/2078-2489/7/1/15

[13] B. Rassouli and D. Gündüz, "Information-theoretic privacy-preserving schemes based on perfect privacy," 2023. [Online]. Available: https://arxiv.org/abs/2301.11754

[14] A. Zamani, T. J. Oechtering, and M. Skoglund, "On the privacy-utility trade-off with and without direct access to the private data," 2022. [Online]. Available: https://arxiv.org/abs/2212.12475

[15] Y. Y. Shkel and H. V. Poor, "A compression perspective on secrecy measures," *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 1, pp. 163–176, 2021.

[16] M. Kocaoglu, A. Dimakis, S. Vishwanath, and B. Hassibi, "Entropic causal inference," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, Feb. 2017. [Online]. Available: https://ojs.aaai.org/index.php/AAAI/article/view/10674

[17] S. Compton, K. Greenewald, D. A. Katz, and M. Kocaoglu, "Entropic causal inference: Graph identifiability," in *Proceedings of the 39th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, and S. Sabato, Eds., vol. 162. PMLR, 17–23 Jul 2022, pp. 4311–4343. [Online]. Available: https://proceedings.mlr.press/v162/compton22a.html

[18] M. Kocaoglu, A. G. Dimakis, S. Vishwanath, and B. Hassibi, "Entropic causality and greedy minimum entropy coupling," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 1465–1469.

[19] A. Painsky, S. Rosset, and M. Feder, "Innovation representation of stochastic processes with application to causal inference," *IEEE Transactions on Information Theory*, vol. 66, no. 2, pp. 1136–1154, 2020.

[20] ——, "Memoryless representation of markov processes," in *2013 IEEE International Symposium on Information Theory*, 2013, pp. 2294–298.

[21] F. Cicalese, L. Gargano, and U. Vaccaro, "Minimum-entropy couplings and their applications," *IEEE Transactions on Information Theory*, vol. 65, no. 6, pp. 3436–3451, 2019.

[22] C. T. Li, "Efficient approximate minimum entropy coupling of multiple probability distributions," *IEEE Transactions on Information Theory*, vol. 67, no. 8, pp. 5259–5268, 2021.

[23] S. Compton, "A tighter approximation guarantee for greedy minimum entropy coupling," in *2022 IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 168–173.

[24] Y. Y. Shkel and A. Kumar Yadav, "Information spectrum converse for minimum entropy couplings and functional representations," in *2023 IEEE International Symposium on Information Theory (ISIT)*, 2023, pp. 66–71.

[25] S. Compton, D. Katz, B. Qi, K. Greenewald, and M. Kocaoglu, "Minimum-entropy coupling approximation guarantees beyond the majorization barrier," in *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, F. Ruiz, J. Dy, and J.-W. van de Meent, Eds., vol. 206. PMLR, 25–27 Apr 2023, pp. 10 445–10 469. [Online]. Available: https://proceedings.mlr.press/v206/compton23a.html