

References

- [1] J. ABERNETHY, P. L. BARTLETT, A. RAKHLIN, AND A. TEWARI, *Optimal strategies and minimax lower bounds for online convex games*, in Proceedings of the 21st Annual Conference on Learning Theory (COLT 2008), Dec. 2008, pp. 415–423.
- [2] P. L. BARTLETT, *Online prediction*, tech. rep., UC Berkeley, 2015. <http://www.stat.berkeley.edu/~bartlett/papers/b-ol-16.pdf>.
- [3] P. L. BARTLETT, P. GRUNWALD, P. HARREMOES, F. HEDAYATI, AND W. KOTLOWSKI, *Horizon-independent optimal prediction with log-loss in exponential families*, in Proceedings of the Conference on Learning Theory (COLT2013), vol. 30, 2013, pp. 639–661.
- [4] P. L. BARTLETT, E. HAZAN, AND A. RAKHLIN, *Adaptive online gradient descent*, in Advances in Neural Information Processing Systems 20, J. Platt, D. Koller, Y. Singer, and S. Roweis, eds., Cambridge, MA, Sept. 2008, MIT Press, pp. 65–72.
- [5] P. L. BARTLETT, W. KOOLEN, A. MALEK, E. TAKIMOTO, AND M. WARMUTH, *Minimax fixed-design linear regression*, in Proceedings of the Conference on Learning Theory (COLT2015), vol. 40, June 2015, pp. 226–239. <http://jmlr.org/proceedings/papers/v40/Bartlett15.pdf>.
- [6] P. L. BARTLETT AND A. MALEK, *Horizon-independent minimax linear regression*, tech. rep., UC Berkeley, 2017.
- [7] L. BREGMAN, *The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming*, USSR Computational Mathematics and Mathematical Physics, 7 (1967), pp. 200 – 217.
- [8] N. CESA-BIANCHI, Y. FREUND, D. HAUSSLER, D. P. HELMBOLD, R. SCHAPIRE, AND M. K. WARMUTH, *How to use expert advice*, Journal of the ACM, 44 (1997), pp. 427–485.
- [9] B. CLARKE AND A. BARRON, *Jeffreys’ prior is asymptotically least favorable under entropy risk*, Journal of Statistical Planning and Inference, 41 (1994), pp. 37 – 60.
- [10] B. S. CLARKE AND A. R. BARRON, *Information-theoretic asymptotics of bayes methods*, IEEE Transactions on Information Theory, 36 (1990), pp. 453–471.
- [11] J. DUCHI, E. HAZAN, AND Y. SINGER, *Adaptive subgradient methods for online learning and stochastic optimization*, JMLR, 12 (2011), pp. 2121–2159.
- [12] E. HAZAN, *Introduction to online convex optimization*, Foundations and Trends in Optimization, 2, pp. 157–325. <http://ocobook.cs.princeton.edu/>.
- [13] F. HEDAYATI AND P. BARTLETT, *The optimality of Jeffreys prior for online density estimation and the asymptotic normality of maximum likelihood estimators*, in Proceedings of the Conference on Learning Theory (COLT2012), vol. 23, June 2012, pp. 7.1–7.13.
- [14] F. HEDAYATI AND P. L. BARTLETT, *Exchangeability characterizes optimality of sequential normalized maximum likelihood and Bayesian prediction with Jeffreys prior*, in Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics (AISTATS), M. Girolami and N. Lawrence, eds., vol. 22, Apr. 2012, pp. 504–510.
- [15] F. HEDAYATI AND P. L. BARTLETT, *Exchangeability characterizes optimality of sequential normalized maximum likelihood and Bayesian prediction*, tech. rep., UC Berkeley, 2015. <http://www.stat.berkeley.edu/~bartlett/papers/hb-ecosnmlbp-15.pdf>.
- [16] A. KALAI AND S. VEMPALA, *Efficient algorithms for online decision problems*, Journal of Computer and System Sciences, 71 (2005), pp. 291 – 307.
- [17] A. KARLIN AND Y. PERES, *Game Theory, Alive*, 2016.

- [18] P. KONTKANEN, W. BUNTINE, P. MYLLYMAKI, J. RISSANEN, AND H. TIRRI, *Efficient computation of stochastic complexity*, in Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics, C. M. Bishop and B. J. Frey, eds., Society for Artificial Intelligence and Statistics, 2003, pp. 181–188.
- [19] P. KONTKANEN AND P. MYLLYMAKI, *A fast normalized maximum likelihood algorithm for multinomial data*, in Proceedings of IJCAI05, 2005.
- [20] W. KOOLEN, A. MALEK, P. L. BARTLETT, AND Y. ABBASI-YADKORI, *Minimax time series prediction*, in Advances in Neural Information Processing Systems 28, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, R. Garnett, and R. Garnett, eds., Curran Associates, Inc., 2015, pp. 2548–2556. <http://papers.nips.cc/paper/5730-minimax-time-series-prediction.pdf>.
- [21] W. M. KOOLEN, A. MALEK, AND P. L. BARTLETT, *Efficient minimax strategies for square loss games*, in Advances in Neural Information Processing Systems 27, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, eds., Curran Associates, Inc., 2014, pp. 3230–3238. <http://papers.nips.cc/paper/5243-efficient-minimax-strategies-for-square-loss-games.pdf>.
- [22] N. LITTLESTONE, *Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm*, Machine Learning, 2 (1988), pp. 285–318.
- [23] N. LITTLESTONE AND M. WARMUTH, *The weighted majority algorithm*, Information and Computation, 108 (1994), pp. 212–261.
- [24] A. S. NEMIROVSKY AND D. B. YUDIN, *Problem Complexity and Method Efficiency in Optimization*, Wiley-Interscience, 1983.
- [25] A. RAKHLIN AND K. SRIDHARAN, *Statistical learning and sequential prediction*, tech. rep., 2014. http://www-stat.wharton.upenn.edu/~rakhlin/courses/stat928/stat928_notes.pdf.
- [26] S. SHALEV-SHWARTZ, *Online learning and online convex optimization*, Foundations and Trends in Machine Learning, 4 (2012).
- [27] E. TAKIMOTO AND M. K. WARMUTH, *The minimax strategy for Gaussian density estimation*, in 13th COLT, 2000, pp. 100–106.
- [28] M. ZINKEVICH, *Online convex programming and generalized infinitesimal gradient ascent*, in Proceedings of the Twentieth International Conference on Machine Learning, 2003, pp. 928–936.