

References

- [1] Stephen Abbott. *Understanding analysis*. Springer, 2001.
- [2] David F Andrews and Colin L Mallows. Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 99–102, 1974.
- [3] Edwin F Beckenbach and Richard Bellman. *Inequalities*, volume 30. Springer Science & Business Media, 2012.
- [4] John Duchi and Yoram Singer. Efficient online and batch learning using forward backward splitting. *Journal of Machine Learning Research*, 10(Dec):2899–2934, 2009.
- [5] Sébastien Gerchinovitz. Sparsity regret bounds for individual sequences in online linear regression. *Journal of Machine Learning Research*, 14(Mar):729–769, 2013.
- [6] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.
- [7] Jyrki Kivinen and Manfred Warmuth. Averaging expert predictions. In *Computational Learning Theory*, pages 638–638. Springer, 1999.
- [8] Jarosław Kotowicz. Convergent real sequences. upper and lower bound of sets of real numbers. *Formalized Mathematics*, 1(3):477–481, 1990.
- [9] John Langford, Lihong Li, and Tong Zhang. Sparse online learning via truncated gradient. *Journal of Machine Learning Research*, 10(Mar):777–801, 2009.
- [10] Francesco Orabona, Nicolo Cesa-Bianchi, and Claudio Gentile. Beyond logarithmic bounds in online learning. In *Artificial Intelligence and Statistics*, pages 823–831, 2012.
- [11] Trevor Park and George Casella. The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008.
- [12] Bala Rajaratnam, Steven Roberts, Doug Sparks, and Onkar Dalal. Lasso regression: estimation and shrinkage via the limit of gibbs sampling. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(1):153–174, 2016.
- [13] Christian Robert. *Machine learning, a probabilistic perspective*. Taylor & Francis, 2014.
- [14] Walter Rudin et al. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1976.
- [15] Rajiv Sambasivan, Sourish Das, and Sujit K Saha. A bayesian perspective of statistical machine learning for big data. *arXiv preprint arXiv:1811.04788*, 2018.
- [16] Shai Shalev-Shwartz and Ambuj Tewari. Stochastic methods for l_1 -regularized loss minimization. *Journal of Machine Learning Research*, 12(Jun):1865–1892, 2011.
- [17] Terence Tao. *An introduction to measure theory*. American Mathematical Society Providence, RI, 2011.
- [18] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.
- [19] Ryan J Tibshirani et al. The lasso problem and uniqueness. *Electronic Journal of Statistics*, 7:1456–1490, 2013.
- [20] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. Technical Report CMU-CS-03-110, School of Computer Science, Carnegie Mellon University, 2003.