Signal Decomposition Using Masked Proximal Operators

Other titles in Foundations and Trends® in Signal Processing

An Introduction to Quantum Machine Learning for Engineers
Osvaldo Simeone
ISBN: 978-1-63828-058-3

Bilevel Methods for Image Reconstruction
Caroline Crockett and Jeffrey A. Fessler
ISBN: 978-1-63828-002-6

Operating Characteristics for Classical and Quantum Binary Hypothesis Testing
Catherine A. Medlock and Alan V. Oppenheim
ISBN: 978-1-68083-882-4

Foundations of User-Centric Cell-Free Massive MIMO
Özlem Tugfe Demir, Emil Björnson and Luca Sanguinetti
ISBN: 978-1-68083-790-2

Data-Driven Multi-Microphone Speaker Localization on Manifolds
Bracha Laufer-Goldshtein, Ronen Talmon and Sharon Gannot
ISBN: 978-1-68083-736-0

Signal Decomposition Using Masked Proximal Operators

Bennet E. Meyers
Stanford University
bennetm@stanford.edu

Stephen P. Boyd
Stanford University
boyd@stanford.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Signal Processing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

B. E. Meyers and S. P. Boyd. *Signal Decomposition Using Masked Proximal Operators*. Foundations and Trends® in Signal Processing, vol. 17, no. 1, pp. 1–78, 2023.

ISBN: 978-1-63828-103-0

Foundations and Trends® in Signal Processing
Volume 17, Issue 1, 2023

Editorial Scope

Topics

Foundations and Trends® in Signal Processing publishes survey and tutorial articles in the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multirate signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
  - Classification and detection
  - Estimation and regression
  - Tree-structured methods

Information for Librarians

Foundations and Trends® in Signal Processing, 2023, Volume 17, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.

Contents

Signal Decomposition Using Masked Proximal Operators

Bennet E. Meyers[1] and Stephen P. Boyd[2]

[1] *Stanford University, USA; bennetm@stanford.edu*
[2] *Stanford University, USA; boyd@stanford.edu*

ABSTRACT

We consider the well-studied problem of decomposing a vector time series signal into components with different characteristics, such as smooth, periodic, nonnegative, or sparse. We describe a simple and general framework in which the components are defined by loss functions (which include constraints), and the signal decomposition is carried out by minimizing the sum of losses of the components (subject to the constraints). When each loss function is the negative log-likelihood of a density for the signal component, this framework coincides with maximum a posteriori probability (MAP) estimation; but it also includes many other interesting cases. Summarizing and clarifying prior results, we give two distributed optimization methods for computing the decomposition, which find the optimal decomposition when the component class loss functions are convex, and are good heuristics when they are not. Both methods require only the masked proximal operator of each of the component loss functions, a generalization of the well-known proximal operator that handles missing entries in its argument. Both methods are distributed, *i.e.*, handle each component separately. We derive tractable methods for evaluating the

Bennet E. Meyers and Stephen P. Boyd (2023), “Signal Decomposition Using Masked Proximal Operators”, Foundations and Trends® in Signal Processing: Vol. 17, No. 1, pp 1–78. DOI: 10.1561/200000122.

masked proximal operators of some loss functions that, to our knowledge, have not appeared in the literature.

1

Introduction

The decomposition of a time series signal into components is an age old problem, with many different approaches proposed, including traditional filtering and smoothing, seasonal-trend decomposition, Fourier and other decompositions, PCA and newer variants such as nonnegative matrix factorization, various statistical methods, and many heuristic methods. It is believed that ancient Babylonian mathematicians used harmonic analysis to understand astronomical observations as collections of 'periodic phenomena' [58].

As we will discuss in detail in §3, formulating the problem of decomposing a time series signal into components as an optimization problem has a long history. We introduce a simple framework that unifies many existing approaches, where components are described by their loss functions. Once the component class loss functions are chosen, we minimize the total loss subject to replicating the given signal with the components. We give a simple unified algorithm, based on variations of well-known algorithms, for carrying out this decomposition, which is guaranteed to find the globally optimal decomposition when the loss functions are all convex, and is a good heuristic when they are not. The method accesses the component loss functions only through a modified proximal operator interface, which takes into account that some data in the original signal may be missing. The method is distributed, in that each component class is handled separately, with the algorithm coordinating them.

Handling of missing data. The methods discussed in this monograph are designed to handle missing data in the original signal to be decomposed, a common situation in many practical settings. The signal components in the decomposition, however, do not have any missing data; by summing the components in the decomposition, we obtain a guess or estimate of the missing values in the original signal. This means that signal decomposition can be used as a sophisticated method for guessing or imputing or interpolating missing or unknown entries in a signal. This allows us to carry out a kind of validation or self-consistency check on a decomposition, by pretending that some known entries are missing, and comparing the imputed values to the known ones.

Expressivity and interpretability. The general framework described here includes many well-known problems as specific instances, and it enables the design of newer, more complex components classes than traditional simple ones such as a periodic signal, a trend, a smooth signal, and so on. For example we can define a signal component class that consists of periodic, smooth, and nonnegative signals, or piecewise constant signals that have no more than some specified number of jumps. The resulting decomposition is always interpretable, since we specify the component classes.

Outline. We describe the signal decomposition framework in §2, where we pose signal decomposition as an optimization problem, concluding with an illustrative simple example in §2.9. In §3 we cover related and previous work and methods. Two distributed methods for solving the signal decomposition problem, based on variations of well established algorithms, are described in §4. The next two sections concern loss functions for signal component classes: general attributes are described in §5 and some example classes in §6. The topic of how to fit component class losses given archetypal examples is discussed in §6.3. We conclude the monograph with examples using real data: Weekly CO2 measurements at Mauna Loa in §7.1, hourly traffic over a New York bridge in §7.2, and 1-minute power output for a group (fleet) of seven photo-voltaic (PV) installations in §7.3.

Software. Our monograph is accompanied by an open-source software implementation called `OSD`, short for 'Optimization(-based) Signal Decomposition', available at:

https://github.com/cvxgrp/signal-decomposition.

References

[1] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems", *Journal of Control and Decision*, vol. 5, no. 1, 2018, pp. 42–60.

[2] D. Amelunxen, M. Lotz, M. B. McCoy, and J. A. Tropp, "Living on the edge: Phase transitions in convex programs with random data", *Information and Inference*, vol. 3, no. 3, 2014, pp. 224–294. DOI: 10.1093/imaiai/iau005.

[3] O. Anderson, "On the logic of the decomposition of statistical series into separate components", *Journal of the Royal Statistical Society*, vol. 90, no. 3, 1927, pp. 548–569. DOI: 10.2307/2341204.

[4] L. Avila and J. Cangialosi, "Tropical cyclone report for hurricane Irene (AL092011)", *NOAA National Hurricane Center*, 2013, URL: https://www.nhc.noaa.gov/data/tcr/AL092011_Irene.pdf.

[5] F. Bach, "Structured sparsity-inducing norms through submodular functions", in *Advances in Neural Information Processing Systems*, J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, Eds., vol. 23, Curran Associates, Inc., 2010. DOI: 10.48550/arXiv.1008.4220.

[6] R. Baillie and S.-K. Chung, "Modeling and forecasting from trend-stationary long memory models with applications to climatology", *International Journal of Forecasting*, vol. 18, no. 2, 2002, pp. 215–226. DOI: 10.1016/S0169-2070(01)00154-6.

[7] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing", *IEEE Transactions on Information Theory*, vol. 56, no. 4, 2010, pp. 1982–2001. DOI: 10.1109/TIT.2010.2040894.

[8] A. Beck and L. Tetruashvili, "On the convergence of block coordinate descent type methods", *SIAM Journal on Optimization*, vol. 23, no. 4, 2013, pp. 2037–2060. DOI: 10.1137/120887679.

[9] D. Bertsekas, *Nonlinear Programming: Third Edition*. Nashua, NH: Athena Scientific, 2016.

[10] M. Best and N. Chakravarti, "Active set algorithms for isotonic regression; a unifying framework", *Mathematical Programming*, vol. 47, 1990, pp. 425–439. DOI: 10.1007/BF01580873.

[11] "February 9–10, 2010 North American blizzard", *Wikipedia*, URL: https://en.wikipedia.org/w/index.php?title=February_9%E2%80%9310,_2010_North_American_blizzard&oldid=1030398339.

[12] "December 26–27th 2010 blizzard", *National Weather Service*, URL: https://www.weather.gov/okx/storm12262010.

[13] "January 26–27 2015 blizzard", *National Weather Service*, URL: https://www.weather.gov/okx/Blizzard_01262715.

[14] "January 31–February 2 2021 winter storm", *National Weather Service*, URL: https://www.weather.gov/okx/WinterStormJan31_Feb22021.

[15] P. Bloomfield, "Trends in global temperature", *Climatic Change*, vol. 21, no. 1, 1992, pp. 1–16. DOI: 10.1007/BF00143250.

[16] P. Bloomfield and D. Nychka, "Climate spectra and detecting climate change", *Climatic Change*, vol. 21, no. 3, 1992, pp. 275–287. DOI: 10.1007/BF00139727.

[17] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation", *IEEE Transactions on Image Processing*, vol. 2, no. 3, 1993, pp. 296–310. DOI: 10.1109/83.236536.

[18] M. Boyd, "High-speed monitoring of multiple grid-connected photovoltaic array configurations", *NIST Technical Note 1896*, 2015. DOI: 10.6028/NIST.TN.1896.

[19] M. Boyd, T. Chen, and B. Doughert, *NIST Campus Photovoltaic (PV) Arrays and Weather Station Data Sets*. National Institute of Standards and Technology, 2017. DOI: 10.18434/M3S67G.

[20] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers", *Foundations and Trends in Machine Learning*, vol. 3, no. 1, 2011, pp. 1–122. DOI: 10.1561/2200000016.

[21] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2009.

[22] S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra*. Cambridge university press, 2018.

[23] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?", *Journal of the ACM*, vol. 58, no. 3, 2011, pp. 1–37. DOI: 10.1145/1970392.1970395.

[24] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization", *Foundations of Computational Mathematics*, vol. 9, no. 6, 2009, pp. 717–772. DOI: 10.1007/s10208-009-9045-5.

[25] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems", *Foundations of Computational Mathematics*, vol. 12, no. 6, 2010, pp. 805–849. DOI: 10.1007/s10208-012-9135-7.

[26] J. F. Claerbout and F. Muir, "Robust modeling with erratic data", *Geophysics*, vol. 38, no. 5, 1973, pp. 826–844. DOI: 10.1190/1.1440378.

[27] R. Cleveland, W. Cleveland, J. McRae, and I. Terpenning, "STL: A seasonal-trend decomposition procedure based on loess", *Journal of Official Statistics*, vol. 6, no. 1, 1990, pp. 3–73, URL: https://www.proquest.com/scholarly-journals/stl-seasonal-trend-decomposition-procedure-based/docview/1266805989/se-2.

[28] P. Combettes and J. Pesquet, "Proximal splitting methods in signal processing", *Springer Optimization and Its Applications*, vol. 49, 2011, pp. 185–212. DOI: 10.1007/978-1-4419-9569-8_10.

[29] A. Cuomo, "Governor Cuomo signs the 'New York State on PAUSE' executive order, March 20, 2020", *New York Governor's Press Office*, URL: https://web.archive.org/web/20200328191630/https://www.governor.ny.gov/news/governor-cuomo-signs-new-york-state-pause-executive-order.

[30] S. Diamond and S. Boyd, "CVXPY: A Python-embedded modeling language for convex optimization", *Journal of Machine Learning Research*, vol. 17, no. 83, 2016, pp. 1–5.

[31] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition", *IEEE Transactions on Information Theory*, vol. 47, no. 7, 2001, pp. 2845–2862. DOI: 10.1109/18.959265.

[32] R. Farebrother, "Adrien-Marie Legendre", in *Statisticians of the Centuries*, C. Heyde, E. Seneta, P. Crépel, S. Fienberg, and J. Gani, Eds., New York, NY: Springer, 2001, pp. 101–104. DOI: 10.1007/978-1-4613-0179-0_20.

[33] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography", *Communications of the ACM*, vol. 24, no. 6, 1981, pp. 381–395. DOI: 10.1145/358669.358692.

[34] E. Franklin, "Solar photovoltaic (PV) system components", *The University of Arizona College of Agriculture & Life Sciences*, 2018, pp. 1–8, URL: https://projects.sare.org/wp-content/uploads/az1742-2017.pdf.

[35] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation", *Computers and Mathematics with Applications*, vol. 2, no. 1, 1976, pp. 17–40. DOI: 10.1016/0898-1221(76)90003-1.

[36] R. Glowinski and A. Marroco, "Sur l'approximation, par elements finis d'ordre un, et la resolution, par penalisation-dualité, d'une classe de problems de Dirichlet non lineares", *Revue Française d'Automatique, Informatique, et Recherche Opérationelle*, vol. 9, no. R-2, 1975, pp. 41–76. DOI: 10.1051/m2an/197509R200411.

[37] S. Greenland and M. Longnecker, "Methods for trend estimation from summarized dose-response data, with applications to meta-analysis", *American Journal of Epidemiology*, vol. 135, no. 11, 1992, pp. 1301–1309. DOI: 10.1093/oxfordjournals.aje.a116237.

[38] S. Grotzinger and C. Witzgall, "Projections onto order simplexes", *Applied Mathematics and Optimization*, vol. 12, no. 1, 1984, pp. 247–270. DOI: 10.1007/BF01449044.

[39] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media, 2013.

[40] R. Hodrick and E. Prescott, "Postwar U.S. business cycles: An empirical investigation", *Journal of Money, Credit and Banking*, vol. 29, no. 1, 1997, pp. 1–16. DOI: 10.2307/2953682.

[41] A. Hoerl and R. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems", *Technometrics*, vol. 12, no. 1, 1970, pp. 55–67. DOI: 10.1080/00401706.1970.10488634.

[42] P. Huber, "Robust estimation of a location parameter", *The Annals of Mathematical Statistics*, vol. 35, no. 1, 1964, pp. 73–101. DOI: 10.1214/aoms/1177703732.

[43] P. Huber, *Robust statistics*, vol. 523. Hoboken, NJ: John Wiley & Sons, 1981. DOI: 10.1002/0471725250.

[44] R. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts: Melbourne, Australia, 2018.

[45] P. Ineichen, "A broadband simplified version of the Solis clear sky model", *Solar Energy*, vol. 82, no. 8, 2008, pp. 758–762. DOI: 10.1016/j.solener.2008.02.009.

[46] R. Inman, Y. Chu, and C. Coimbra, "Cloud enhancement of global horizontal irradiance in California and Hawaii", *Solar Energy*, vol. 130, 2016, pp. 128–138. DOI: 10.1016/j.solener.2016.02.011.

[47] S.-J. Kim, K. Koh, S. Boyd, and D. Gorinevsky, "L1 trend filtering", *SIAM Review*, vol. 51, no. 2, 2009, pp. 339–360. DOI: 10.1137/070690274.

[48] R. Koenker and G. Bassett, "Regression quantiles", *Econometrica*, vol. 46, no. 1, 1978, p. 33. DOI: 10.2307/1913643.

[49] R. Koenker and K. F. Hallock, "Quantile regression", *Journal of Economic Perspectives*, vol. 15, no. 4, 2001, pp. 143–156. DOI: 10.1257/jep.15.4.143.

[50] C. Leser, "A simple method of trend construction", *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 23, no. 1, 1961, pp. 91–107. DOI: 10.1111/j.2517-6161.1961.tb00393.x.

[51] S. Levitt, “Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not”, *Journal of Economic Perspectives*, vol. 18, no. 1, 2004, pp. 163–190. DOI: 10.1257/089533004773563485.

[52] W. Link and F. Sauer, “Estimating equations estimates of trends”, *Bird Populations*, vol. 2, 1994, pp. 23–32, URL: https://bmeyers.github.io/assets/link-1994.pdf.

[53] S. Mallat, *A Wavelet Tour of Signal Processing*. Elsevier, 2009, pp. 20–41. DOI: 10.1016/B978-0-12-374370-1.X0001-8.

[54] M. B. McCoy, V. Cevher, Q. T. Dinh, A. Asaei, and L. Baldassarre, “Convexity in source separation: Models, geometry, and algorithms”, *IEEE Signal Processing Magazine*, vol. 31, no. 3, 2014, pp. 87–95. DOI: 10.1109/MSP.2013.2296605.

[55] M. B. McCoy and J. A. Tropp, “Sharp recovery bounds for convex demixing, with applications”, *Foundations of Computational Mathematics*, vol. 14, no. 3, 2014, pp. 503–567. DOI: 10.1007/s10208-014-9191-2.

[56] J.-J. Moreau, “Fonctions convexes duales et points proximaux dans un espace hilbertien”, *Reports of the Paris Academy of Sciences, Series A*, vol. 255, 1962, pp. 2897–2899.

[57] “Hourly traffic on Metropolitan Transportation Authority (MTA) bridges and tunnels”, *NY Open Data*, URL: https://data.ny.gov/Transportation/Hourly-Traffic-on-Metropolitan-Transportation-Auth/qzve-kjga.

[58] O. Neugebauer, *The Exact Sciences in Antiquity*, ser. Acta historica scientiarum naturalium et medicinalium. Dover Publications, 1969, URL: https://books.google.com/books?id=JVhTtVA2zr8C.

[59] A. Oppenheim and R. Schafer, *Discrete-time Signal Processing*, ser. Prentice-Hall signal processing series. United Kingdom: Pearson, 2010.

[60] D. Osborn, “Moving average detrending and the analysis of business cycles”, *Oxford Bulletin of Economics and Statistics*, vol. 57, no. 4, 1995, pp. 547–558. DOI: 10.1111/j.1468-0084.1995.tb00039.x.

[61] N. Parikh and S. Boyd, “Proximal algorithms”, *Foundations and Trends in Optimization*, vol. 1, no. 3, 2014, pp. 127–239. DOI: 10.1561/2400000003.

[62] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python", *Journal of Machine Learning Research*, vol. 12, 2011, pp. 2825–2830, URL: https://jmlr.org/papers/v12/pedregosa11a.html.

[63] D. Phillips, "A technique for the numerical solution of certain integral equations of the first kind", *Journal of the ACM*, vol. 9, no. 1, 1962, pp. 84–97. DOI: 10.1145/321105.321114.

[64] R. Poliquin and R. Rockafellar, "Prox-regular functions in variational analysis", *Transactions of the American Mathematical Society*, vol. 348, no. 5, 1996, pp. 1805–1838. DOI: 10.1090/S0002-9947-96-01544-9.

[65] P. Richtárik and M. Takáč, "Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function", *Mathematical Programming*, vol. 144, no. 1-2, 2014, pp. 1–38. DOI: 10.1007/s10107-012-0614-z.

[66] R. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

[67] R. Sah, "Caltrans traffic census program", *California Department of Transportation*, URL: https://dot.ca.gov/programs/traffic-operations/census.

[68] P. Sen, "Estimates of the regression coefficient based on Kendall's Tau", *Journal of the American Statistical Association*, vol. 63, no. 324, 1968, pp. 1379–1389. DOI: 10.1080/01621459.1968.10480934.

[69] K. Singleton, "Econometric issues in the analysis of equilibrium business cycle models", *Journal of Monetary Economics*, vol. 21, no. 2-3, 1988, pp. 361–386. DOI: 10.1016/0304-3932(88)90036-0.

[70] J.-L. Starck, F. Murtagh, and J. M. Fadili, *Sparse image and signal processing: wavelets, curvelets, morphological diversity*. Cambridge, UK: Cambridge university press, 2010. DOI: 10.1017/CBO9780511730344.

[71] J. Perktold, S. Seabold, and J. Taylor, "Seasonal-trend decomposition using LOESS (STL)", *Statsmodels documentation*, URL: https://www.statsmodels.org/dev/examples/notebooks/generated/stl_decomposition.html.

[72] "STL: Seasonal decomposition of time series by loess", *R documentation*, URL: https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/stl.

[73] "Time series decomposition", *MathWorks documentation*, URL: https://www.mathworks.com/help/econ/detrending.html.

[74] P. Tans and R. Keeling, "Mauna Loa CO2 weekly mean and historical comparisons", *NOAA Global Monitoring Laboratory, Earth System Research Laboratories*, URL: https://gml.noaa.gov/ccgg/trends/data.html.

[75] H. L. Taylor, S. C. Banks, and J. F. McCoy, "Deconvolution with the l1 norm", *Geophysics*, vol. 44, no. 1, 1979, pp. 39–52. DOI: 10.1190/1.1440921.

[76] H. Theil, "A rank-invariant method of linear and polynomial regression analysis", *Proceedings of the Royal Netherlands Academy of Sciences*, vol. 53, 1950, Part I: 386–392, Part II: 521–525, Part III: 1397–1412. DOI: 10.1007/978-94-011-2546-8_20.

[77] A. Thompson and J. Kay, "On some Bayesian choices of regularization parameter in image restoration", *Inverse Problems*, vol. 9, no. 6, 1993, pp. 749–761. DOI: 10.1088/0266-5611/9/6/011.

[78] R. Tibshirani, "Regression shrinkage and selection via the lasso", *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, 1996, pp. 267–288. DOI: 10.1111/j.2517-6161.1996.tb02080.x.

[79] A. Tikonov, "Solution of incorrectly formulated problems and the regularization method", *Soviet Math.*, vol. 4, 1963, pp. 1035–1038.

[80] D. Titterington, "General structure of regularization procedures in image reconstruction", *Astronomy and Astrophysics*, vol. 144, no. 2, 1985, pp. 381–387.

[81] I. Tošić and P. Frossard, "Dictionary learning", *IEEE Signal Processing Magazine*, vol. 28, no. 2, 2011, pp. 27–38. DOI: 10.1109/MSP.2010.939537.

[82] R. Tsay, *Analysis of Financial Time Series*, ser. Wiley Series in Probability and Statistics. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2005, pp. 206–215. DOI: 10.1002/0471746193.

[83] M. Udell, C. Horn, R. Zadeh, and S. Boyd, "Generalized low rank models", *Foundations and Trends in Machine Learning*, vol. 9, no. 1, 2016, pp. 1–118. DOI: 10.1561/2200000055.

[84] J. Wright and Y. Ma, *High-dimensional data analysis with low-dimensional models: Principles, computation, and applications*. New York, NY: Cambridge University Press, 2022.

[85] S. Wright, "Coordinate descent algorithms", *Mathematical Programming*, vol. 151, no. 1, 2015, pp. 3–34. DOI: 10.1007/s10107-015-0892-3.

[86] W. B. Wu, M. Woodroofe, and G. Mentz, "Isotonic regression: Another look at the changepoint problem", *Biometrika*, vol. 88, no. 3, 2001, pp. 793–804. DOI: 10.1093/biomet/88.3.793.

[87] M. Wytock and J. Kolter, "Contextually supervised source separation with application to energy disaggregation", *Twenty-Eighth AAAI Conference on Artificial Intelligence*, vol. 28, no. 1, 2013, pp. 486–492. DOI: 10.1609/aaai.v28i1.8769.

[88] A. Yurtsever, V. M., and S. Sra, "Three operator splitting with a nonconvex loss function", 2021. DOI: 10.48550/arXiv.2103.04568.

[89] K. Zhao, M. A. Wulder, T. Hu, R. Bright, Q. Wu, H. Qin, Y. Li, E. Toman, B. Mallick, X. Zhang, and M. Brown, "Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm", *Remote Sensing of Environment*, vol. 232, 2019, pp. 111–181. DOI: 10.1016/j.rse.2019.04.034.

[90] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net", *Journal of the royal statistical society, series B (statistical methodology)*, vol. 67, no. 2, 2005, pp. 301–320. DOI: 10.1111/j.1467-9868.2005.00503.x.