

Le Thi, H. A.; Pham Dinh, T.; Le, H. M.; Vo, X. T.

DC approximation approaches for sparse optimization. (English) Zbl 1346.90819
Eur. J. Oper. Res. 244, No. 1, 26-46 (2015).

Summary: Sparse optimization refers to an optimization problem involving the zero-norm in objective or constraints. In this paper, nonconvex approximation approaches for sparse optimization have been studied with a unifying point of view in DC (Difference of Convex functions) programming framework. Considering a common DC approximation of the zero-norm including all standard sparse inducing penalty functions, we studied the consistency between global minimums (resp. local minimums) of approximate and original problems. We showed that, in several cases, some global minimizers (resp. local minimizers) of the approximate problem are also those of the original problem. Using exact penalty techniques in DC programming, we proved stronger results for some particular approximations, namely, the approximate problem, with suitable parameters, is equivalent to the original problem. The efficiency of several sparse inducing penalty functions have been fully analyzed. Four DCA (DC Algorithm) schemes were developed that cover all standard algorithms in nonconvex sparse approximation approaches as special versions. They can be viewed as, an ℓ_1 -perturbed algorithm/reweighted- ℓ_1 algorithm/reweighted- ℓ_2 algorithm. We offer a unifying nonconvex approximation approach, with solid theoretical tools as well as efficient algorithms based on DC programming and DCA, to tackle the zero-norm and sparse optimization. As an application, we implemented our methods for the feature selection in SVM (Support Vector Machine) problem and performed empirical comparative numerical experiments on the proposed algorithms with various approximation functions.

MSC:

[90C56](#) Derivative-free methods and methods using generalized derivatives

Cited in **31** Documents

[65K05](#) Numerical mathematical programming methods

Keywords:

global optimization; sparse optimization; DC approximation function; DC programming and DCA; feature selection in SVM

Full Text: [DOI](#)

References:

- [1] Bajwa, W.; Haupt, J.; Sayeed, A.; Nowak, R., Compressive wireless sensing, Proceedings of fifth international conference on information processing in sensor networks, 134-142, (2006)
- [2] Baron, D.; Wakin, M. B.; Duarte, M. F.; Sarvotham, S.; Baraniuk, R. G., Distributed compressed sensing, (2009, November), Electrical and Computer Engineering Department, Rice University, Technical report ECE06-12
- [3] Bennett, K. P.; Mangasarian, O. L., Robust linear programming discrimination of two linearly inseparable sets, Optimization Methods and Software, 1, 23-34, (1992)
- [4] Bradley, P. S.; Mangasarian, O. L., Feature selection via concave minimization and support vector machines, Proceeding of international conference on machine learning ICML'98, (1998)
- [5] Bradley, P. S.; Mangasarian, O. L.; Rosen, J. B., Parsimonious least norm approximation, Computational Optimization and Applications, 11, 1, 5-21, (1998) · [Zbl 0979.65032](#)
- [6] Candes, E. J.; Tao, T., Decoding by linear programming, IEEE Transactions on Information Theory, 51, 12, 4203-4215, (2005) · [Zbl 1264.94121](#)
- [7] Candes, E. J.; Randall, P., Highly robust error correction by convex programming, IEEE Transactions on Information Theory, 54, 2829-2840, (2006) · [Zbl 1332.94096](#)
- [8] Candes, E. J.; Wakin, M.; Boyd, S., Enhancing sparsity by reweighted- ℓ_1 minimization, Journal of Mathematical Analysis and Applications, 14, 877-905, (2008) · [Zbl 1176.94014](#)
- [9] Chartrand, R.; Yin, W., Iteratively reweighted algorithms for compressive sensing, ICASSP, (2008)
- [10] Chan, A. B.; Vasconcelos, N.; Lanckriet, R. G., Direct convex relaxations of sparse SVM, Proceeding ICML'07. Proceedings of the 24th international conference on machine learning, 145-153, (2007)
- [11] Chen, X.; Xu, F. M.; Ye, Y., Lower bound theory of nonzero entries in solutions of ℓ_2 -lp minimization, SIAM Journal on

- Scientific Computing, 32, 5, 2832-2852, (2010) · [Zbl 1242.90174](#)
- [12] Cortes, C.; Vapnik, V. N., Support vector networks, *Machine Learning*, 20, 3, 273-297, (1995) · [Zbl 0831.68098](#)
- [13] Collobert, R.; Sinz, F.; Weston, J.; Bottou, L., Trading convexity for scalability, *Proceedings of the 23th international conference on machine learning (ICML 2006)*, (2006), Pittsburgh, PA
- [14] Fan, J.; Li, R., Variable selection via nonconcave penalized likelihood and its oracle properties, *Journal of the American Statistical Association*, 96, 456, 1348-1360, (2001) · [Zbl 1073.62547](#)
- [15] Fawzi, A.; Davies, M.; Frossard, P., Dictionary learning for fast classification based on soft-thresholding, *International Journal of Computer Vision*, (2014), <http://arxiv.org/abs/1402.1973>
- [16] Fu, W. J., Penalized regression: the bridge versus the lasso, *Journal of Computational and Graphical Statistics*, 7, 397-416, (1998)
- [17] Gasso, G.; Rakotomamonjy, A.; Canu, S., Recovering sparse signals with a certain family of nonconvex penalties and dc programming, *IEEE Transactions on Signal Processing*, 57, 4686-4698, (2009) · [Zbl 1391.90489](#)
- [18] Gribonval, R.; Nielsen, M., Sparse representation in union of bases, *IEEE Transactions on Information Theory*, 49, 3320-3325, (2003) · [Zbl 1286.94032](#)
- [19] Gorodnitsky, I. F.; Rao, B. D., Sparse signal reconstructions from limited data using FOCUSS: A re-weighted minimum norm algorithm, *IEEE Transactions on Signal Processing*, 45, 600-616, (1997)
- [20] Guyon, I.; Weston, J.; Barnhill, S.; Vapnik, V. N., Gene selection for cancer classification using support vector machines, *Machine Learning*, 46, 1-3, 389-422, (2002) · [Zbl 0998.68111](#)
- [21] Guyon, I.; Gunn, S.; Nikravesh, M.; Zadeh, L. A., *Feature extraction, foundations and applications*, (2006), Springer Berlin
- [22] Guan, W.; Gray, A., Sparse high-dimensional fractional-norm support vector machine via DC programming, *Computational Statistics and Data Analysis*, 67, 136-148, (2013) · [Zbl 06970878](#)
- [23] Hastie, T.; Tibshirani, R.; Friedman, J., *The elements of statistical learning*, (2009), Springer Heidelberg
- [24] Heiler, M.; Cremers, D.; Schnörr, C., Efficient feature subset selection for support vector machines, (2001), Department of Mathematics and Computer Science, University of Mannheim
- [25] Krause, N.; Singer, Y., Leveraging the margin more carefully, *Proceedings of the 21st international conference on machine learning, ICML 2004*, (2004), Banff Alberta, Canada
- [26] Peleg, D.; Meir, R., A bilinear formulation for vector sparsity optimization, *Signal Processing*, 8, 2, 375-389, (2008) · [Zbl 1186.94273](#)
- [27] Le Thi, H. A., *Analyse numérique des algorithmes de l'Optimisation d. c. Approches locales et globales. Codes et simulations numériques en grande dimension. Applications*, (1994), Université de Rouen, (Thèse de doctorat)
- [28] Le Thi, H. A., *Contribution à l'optimisation non convexe et l'optimisation globale: Théorie, algorithmes et applications*, (1997), Habilitation à Diriger des Recherches, Université de Rouen
- [29] Le Thi, H. A.; Pham Dinh, T., Solving a class of linearly constrained indefinite quadratic problems by DC algorithms, *Journal of Global Optimization*, 11, 3, 253-285, (1997) · [Zbl 0905.90131](#)
- [30] Le Thi, H. A., An efficient algorithm for globally minimizing a quadratic function under convex quadratic constraints, *Mathematical Programming*, 87, 3, 401-426, (2000) · [Zbl 0952.90031](#)
- [31] Le Thi, H. A. (2012). DC programming and DCA. <http://lita.sciences.univ-metz.fr/~lethi> · [Zbl 1261.90082](#)
- [32] Le Thi, H. A.; Pham Dinh, T.; Nguyen Van, T., Combination between local and global methods for solving an optimization problem over the efficient set, *European Journal of Operational Research*, 142, 258-270, (2002) · [Zbl 1082.90563](#)
- [33] Le Thi, H. A.; Pham Dinh, T., Large scale molecular optimization from distance matrices by a D.C. optimization approach, *SIAM Journal on Optimization*, 4, 1, 77-116, (2003) · [Zbl 1075.90071](#)
- [34] Le Thi, H. A.; Pham Dinh, T., The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems, *Annals of Operations Research*, 133, 23-46, (2005) · [Zbl 1116.90122](#)
- [35] Le Thi, H. A.; Belghiti, T.; Pham Dinh, T., A new efficient algorithm based on DC programming and DCA for clustering, *Journal of Global Optimization*, 37, 593-608, (2006) · [Zbl 1198.90327](#)
- [36] Le Thi, H. A.; Le, H. M.; Pham Dinh, T., Optimization based DC programming and DCA for hierarchical clustering, *European Journal of Operational Research*, 183, 1067-1085, (2006) · [Zbl 1149.90117](#)
- [37] Le Thi, H. A.; Nguyen, T. P.; Pham Dinh, T., A continuous approach for solving the concave cost supply problem by combining DCA and B&B techniques, *European Journal of Operational Research*, 183, 1001-1012, (2007)
- [38] Le Thi, H. A.; Le, H. M.; Pham Dinh, T., Optimization based DC programming and DCA for hierarchical clustering, *European Journal of Operational Research*, 183, 1067-1085, (2007) · [Zbl 1149.90117](#)
- [39] Le Thi, H. A.; Pham Dinh, T., A continuous approach for the concave cost supply problem via DC programming and DCA, *Discrete Applied Mathematics*, 156, 325-338, (2008) · [Zbl 1190.90142](#)
- [40] Le Thi, H. A.; Le, H. M.; Nguyen, V. V.; Pham Dinh, T., A dc programming approach for feature selection in support vector machines learning, *Journal of Advances in Data Analysis and Classification*, 2, 259-278, (2008) · [Zbl 1284.90057](#)
- [41] Le Thi, H. A.; Nguyen, V. V.; Ouchani, S., Gene selection for cancer classification using DCA, *Journal of Fonctiers of Computer Science and Technology*, 3, 6, 62, (2009)
- [42] Le Thi, H. A.; Huynh, V. N.; Pham Dinh, T., Convergence analysis of DC algorithms for DC programming with subanalytic data, (2009), National Institute for Applied Sciences Rouen, http://www.optimization-online.org/DB_HTML/2013/08/3996.html

- [43] Le Thi, H. A., A new approximation for the ℓ_0 -norm, (2012), University of Lorraine, (Research report LITA EA 3097)
- [44] Le Thi, H. A.; Huynh, V. N.; Pham Dinh, T., Exact penalty and error bounds in DC programming, *Journal of Global Optimization*, 52, 3, 509-535, (2012) · [Zbl 1242.49037](#)
- [45] Le Thi, H. A.; Moeini, M.; Pham Dinh, T.; Joaquim, J., A DC programming approach for solving the symmetric eigenvalue complementarity problem, *Computational Optimization and Applications*, 51, 3, 1097-1117, (2012) · [Zbl 1241.90153](#)
- [46] Le, H. M.; Le Thi, H. A.; Pham Dinh, T.; Huynh, V. N., Block clustering based on difference of convex functions (DC) programming and DC algorithms, *Neural Computation*, 25, 10, 2776-2807, (2013) · [Zbl 1418.62253](#)
- [47] Le, H. M.; Le Thi, H. A.; Nguyen, M. C., DCA based algorithms for feature selection in semi-supervised support vector machines, (Perner, P., *Machine learning and data mining in pattern recognition*, (2013)), 528-542, LNAI 7988
- [48] Le Thi, H. A.; Le, H. M.; Pham Dinh, T.; Huynh, V. N., Binary classification via spherical separator by DC programming and DCA, *Journal of Global Optimization*, 56, 4, 1393-1407, (2013) · [Zbl 1275.90093](#)
- [49] Le Thi, H. A.; Nguyen, T. B.T.; Le, H. M., Sparse signal recovery by difference of convex functions algorithms, *Lecture notes in computer science*, 387-397, (2013), ISBN 978-3-642-36542-3
- [50] Le Thi, H. A.; Pham Dinh, T., DC programming approaches for distance geometry problems, (Mucherino, A.; Lavor, C.; Liberti, L.; Maculan, N., *Distance geometry: Theory, methods and applications*, (2013), Springer), 225-290 · [Zbl 1271.65038](#)
- [51] Le Thi, H. A.; Nguyen, M. C., Efficient algorithms for feature selection in multi-class support vector machine, *Advanced computational methods for knowledge engineering. Studies in computational intelligence* 479, (2013), Springer
- [52] Le Thi, H. A.; Vo, X. T.; Pham Dinh, T., Feature selection for linear SVMs under uncertain data: robust optimization based on difference of convex functions algorithms., *Neural Networks*, 59, 36-50, (2014) · [Zbl 1327.90236](#)
- [53] Le Thi, H. A.; Nguyen, M. C., Self-organizing maps by difference of convex functions optimization, *Data Mining and Knowledge Discovery*, 28, 5-6, 1336-1365, (2014) · [Zbl 1342.90147](#)
- [54] Le Thi, H. A.; Le, H. M.; Pham Dinh, T., Feature selection in machine learning: an exact penalty approach using a difference of convex function algorithm, *Machine Learning*, (2014), (published online 04.07.14). doi: 10.1007/s10994-014-5455-y
- [55] Le Thi, H. A.; Moeini, M., Long-short portfolio optimization under cardinality constraints by difference of convex functions algorithm, *Journal of Optimization Theory, & Applications*, 161, 1, 199-224, (2014) · [Zbl 1300.91046](#)
- [56] Le Thi, H. A.; Huynh, V. N.; Pham Dinh, T., DC programming and DCA for solving general DC programs, *Proceedings of 2nd international conference on computer science, applied mathematics and applications (ICCSAMA 2014), Advances in Intelligent Systems and Computing*, 282, 15-35, (2014) · [Zbl 1322.90072](#)
- [57] Liu, Y.; Zheng, Y. F., FS SFS: A novel feature selection method for support vector machines, *Pattern Recognition*, (2005)
- [58] Liu, Y.; Shen, X., Multicategory ψ -learning, *Journal of the American Statistical Association*, 101, 500-509, (2006) · [Zbl 1119.62341](#)
- [59] Mallat, S.; Zhang, Z., Matching pursuit in a time-frequency dictionary, *IEEE Transactions on Signal Processing*, 41, 12, 3397-3415, (1993) · [Zbl 0842.94004](#)
- [60] Mangasarian, O. L., *Machine learning via polyhedral concave minimization*, (Fischer, H.; Riedmueller, B.; Schaeffler, S., *Applied mathematics and parallel computing - Festschrift for Klaus Ritter*, (1996), Physica-Verlag Germany), 175-188 · [Zbl 0906.68127](#)
- [61] Mohri, M.; Medina, A. M., Learning theory and algorithms for revenue optimization in second-price auctions with reserve, *Proceeding ICML'14. Proceedings of the 31th international conference on machine learning*, (2014), <http://arxiv.org/abs/1310.5665>
- [62] Natarajan, B. K., Sparse approximate solutions to linear systems, *SIAM Journal on Scientific Computing*, 24, 227-234, (1995) · [Zbl 0827.68054](#)
- [63] Neumann, J.; Schnörr, G.; Steidl, G., Combined SVM-based feature selection and classification., *Machine Learning*, 61, 1-3, 129-150, (2005) · [Zbl 1137.90643](#)
- [64] Niu, Y. S.; Pham Dinh, T.; Le Thi, H. A. Judice, J. (2012). Efficient DC programming approaches for asymmetric eigenvalue complementarity problem, *optimization methods and software*, doi: 10.1080/10556788.2011.645543.
- [65] Ong, C. S.; Le Thi, H. A., Learning sparse classifiers with difference of convex functions algorithms, *Optimization Methods and Software*, 28, 4, 830-854, (2013) · [Zbl 1282.90181](#)
- [66] Pati, Y. C.; Rezaifar, R.; Krishnaprasa, P. S., Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, *27th asilomar conference on signals, systems and computers*, (1993, November)
- [67] Pham Dinh, T.; Le Thi, H. A., Convex analysis approach to d.c. programming: theory, algorithm and applications, *Acta Mathematica Vietnamica*, 22, 289-355, (1997) · [Zbl 0895.90152](#)
- [68] Pham Dinh, T.; Le Thi, H. A., DC optimization algorithms for solving the trust region subproblem., *SIAM Journal of Optimization*, 8, 476-505, (1998) · [Zbl 0913.65054](#)
- [69] Pham Dinh, T.; Le Thi, H. A., DC programming: theory, algorithms and applications. the state of the art, *Proceedings of the first international workshop on global constrained optimization and constraint satisfaction (Cocos' 02)*, (2002), Valbonne-Sophia Antipolis France, (28 pp., October 2-4)
- [70] Pham Dinh, T.; Nguyen Canh, N.; Le Thi, H. A., An efficient combination of DCA and B&B using DC/SDP relaxation for globally solving binary quadratic programs, *Journal of Global Optimization*, 48, 4, 595-632, (2010) · [Zbl 1226.90060](#)
- [71] Pham Dinh, T.; Le Thi, H. A., Recent advances in DC programming and DCA, *Transactions on Computational Collective Intelligence*, 8342, 1-37, (2014)
- [72] Pham Dinh, T.; Le, H. M.; Le Thi, H. A.; Lauer, F., A DC programming algorithm for switched linear regression, *IEEE*

- Transactions on Automatic Control, 59, 8, 2277-2282, (2014) · [Zbl 1360.62386](#)
- [73] Rao, B. D.; Kreutz-Delgado, K., An affine scaling methodology for best basis selection, *IEEE Transactions on Signal Processing*, 47, 87-200, (1999) · [Zbl 0984.94010](#)
- [74] Rao, B. D.; Engan, K.; Cotter, S. F.; Palmer, J.; KreutzDelgado, K., Subset selection in noise based on diversity measure minimization, *IEEE Transactions on Signal Processing*, 51, 3, 760-770, (2003)
- [75] Rinaldi, F.; Schoen, F.; Sciandrone, M., Concave programming for minimizing the zero-norm over polyhedral sets, *Computational Optimization and Applications*, 46, 3, 467-486, (2010) · [Zbl 1229.90170](#)
- [76] Rockafellar, R. T., *Convex analysis*, (1970), Princeton University Princeton · [Zbl 0202.14303](#)
- [77] Schmidt, M.; Fung, G.; Rosales, G., Fast optimization methods for L1 regularization: A comparative study and two new approaches, *ECML 2007, Lecture Notes in Computer Science*, \textit{Proceedings of machine learning}, 4701, 286-297, (2007)
- [78] Sriperumbudur, B. K.; Torres, D. A.; Lanckriet, R. G., Sparse eigen methods by D.C. programming, *Proceeding ICML '07. Proceedings of the 24th international conference on machine learning*, 831-838, (2007)
- [79] Takhar, D.; Laska, J. N.; Wakin, M. B.; Duarte, M. F.; Baron, D.; Sarvotham, S., A new compressive imaging camera architecture using optical-domain compression, *Computational imaging IV at IS&T/SPIE electronic imaging*, (2006, January), San Jose, California
- [80] Tan, M.; Wang, L.; Tsang, I. W., Learning sparse svm for feature selection on very high dimensional datasets., *ICML 2010*, (2010)
- [81] Thiao, M.; Pham Dinh, T.; Le Thi, H. A., DC programming approach for solving a class of nonconvex programs dealing with zero-norm, *Modelling, computation and optimization in information systems and management science*, CCIS 14, 348-357, (2008), Springer-Verlag · [Zbl 1160.90626](#)
- [82] Thiao, M.; Pham Dinh, T.; Le Thi, H. A., A DC programming approach for sparse eigenvalue problem., *Proceedings of the 27th international conference on machine learning*, *ICML 2010*, 1063-1070, (2010)
- [83] Tibshirani, R., Regression shrinkage and selection via the lasso, *Journal of The Royal Statistical Society*, 46, 431-439, (1996) · [Zbl 0850.62538](#)
- [84] Trombettoni, G.; Araya, I.; Neveu, B.; Chabert, G., Inner regions and interval linearizations for global optimization, *Proceedings of the twenty-fifth AAAI conference on artificial intelligence*, 99-104, (2011)
- [85] Weston, J.; Mukherjee, S.; Chapelle, O.; Pontil, M.; Poggio, T.; Vapnik, V., *Feature selection for svms*, *Neural information processing systems*, (2001), MIT Press Cambridge, MA
- [86] Weston, J.; Elisseeff, Scholkopf, A. B.; Tipping, M., Use of the zero-norm with linear models and kernel methods, *Journal of Machine Learning Research*, 3, 1439-1461, (2003) · [Zbl 1102.68605](#)
- [87] Zhang, T., Some sharp performance bounds for least squares regression with regularization, *Annals of Statistics*, 37, 2109-2144, (2009) · [Zbl 1173.62029](#)
- [88] Zhang, H. H.; Ahn, J.; Lin, X.; Park, C., Gene selection using support vector machines with non-convex penalty, *Bioinformatics*, 2, 1, 88-95, (2006)
- [89] Zou, H., The adaptive lasso and its oracle properties, *Journal of the American Statistical Association*, 101, 1418-1429, (2006) · [Zbl 1171.62326](#)
- [90] Zou, H.; Li, R., One-step sparse estimates in nonconcave penalized likelihood models, *Annals of Statistics*, 36, 4, 1509-1533, (2008) · [Zbl 1142.62027](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.