SCOPT: Self-Concordant OPTimization

SCOPT is a MATLAB implementation of the proximal gradient, proximal quasi-Newton, proximal Newton, and path-following interior-point algorithms for solving composite convex minimization problems involving self-concordant and self-concordant-like functions:

$$\min_{\mathbf{x}\in\mathbb{R}^n}\Big\{F(\mathbf{x}) := f(\mathbf{x}) + g(\mathbf{x})\Big\},$$

where f is a self-concordant or self-concordant-like function, and g is a proper, closed and convex function.

Briefly, the function f is called self-concordant (with the parameter $M_f > 0$) if $|\varphi'''(t)| \leq M_f\varphi''(t)^{3/2}$, where $\varphi(t) := f(\mathbf{x} + t\mathbf{u})$ for $\mathbf{x} \in \mathrm{dom}(f), \mathbf{u} \in \mathbb{R}^m$, and $\mathbf{x} + t\mathbf{u} \in \mathrm{dom}(f)$. Examples: $f(\mathbf{X}) := -\log\det(\mathbf{X})$, $f(\mathbf{x}, t) := -\log(t^2 - \|\mathbf{x}\|_2^2)$, and $f(\mathbf{x}) = -\sum_{i=1}^{p}\log(\mathbf{a}_i^T\mathbf{x} - b_i)$ are all self-concordant.

The function g is usually assumed that its proximal operator

$$\mathrm{prox}_g(\mathbf{x}) := \arg\min_{\mathbf{u}\in\mathbb{R}^n}\big\{g(\mathbf{u}) + (1/2)\|\mathbf{u} - \mathbf{x}\|_2^2\big\}$$

is “easy” to compute (e.g., in a closed form or in polynomial time). SCOPT consists of several algorithms customized for solving the following convex optimization problems (but not limited to those):

- **Sparse Inverse Covariance Estimation Problems**

J. Friedman, T. Hastie, and R. Tibshirani. Sparse Inverse Covariance Estimation with the graphical LASSO. Biostatistics, vol. 9, no. 3, pp. 432?441, 2007.

Example:

$$\min_{\mathbf{X}\succeq 0}\Big\{-\log\det(\mathbf{X}) + \mathrm{tr}(\mathbf{S}\mathbf{X}) + \rho\|\mathrm{vec}(\mathbf{X})\|_1\Big\}.$$

- **Poisson Intensity Reconstruction Problems**

Z. T. Harmany, R. F. Marcia, and R. M. Willett. This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction Algorithms - Theory and Practice. IEEE Trans. Image Processing, vol. 21, no. 3, pp. 1084-1096, 2012.

Example:

$$\min_{\mathbf{x}\geq 0}\Big\{\sum_{i=1}^{m}\mathbf{a}_i^T\mathbf{x} - \sum_{i=1}^{m} b_i\log(\mathbf{a}_i^T\mathbf{x} + \mu_i) + \rho\mathrm{TV}(\mathbf{x})\Big\}.$$

- **Heteroscedastic LASSO problems**

N. Stadler, P. Buhlmann, and S. V. de Geer. L1-Penalization for mixture regression models. TEST, vol. 19, no. 2, pp. 209-256, 2010.

Example:

$$\min_{\sigma>0,\beta\in\mathbb{R}^p}\Big\{-\log(\sigma) + (1/(2n))\|\mathbf{X}\beta - \sigma\mathbf{y}\|_2^2 + \rho\|\beta\|_1\Big\}.$$

- **Sparse logistic and multinomial logistic regression problems**

B. Krishnapuram, M. Figueredo, L. Carin, and A. Hertemink. Sparse multinomial logistic regression: Fast algorithms and generalization bounds. IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), vol. 27, pp. 957-968, 2005.

Example:

$$\min_{\mathbf{X}\in\mathbb{R}^{m\times p}} \Big\{(1/N)\sum_{i=1}^{N}\Big[\log\big(1+\sum_{j=1}^{m}\exp((\mathbf{W}^{(j)})^T\mathbf{X}^{(i)})\big)-\sum_{i=1}^{m}y_i^{(j)}(\mathbf{W}^{(j)})^T\mathbf{X}^{(i)}\Big]+\rho\|\mathrm{vec}(\mathbf{X})\|_1\Big\},$$

where $y^{(j)}$ and $\mathbf{W}^{(j)}$ are input data, $j=1,\ldots,N$.