Deep Learning and Inverse Problems
Summer 2024
Lecturer: Reinhard Heckel

Problem Set 10

Issued: Tuesday July 2, 2024, 1:00 pm.
Due: Tuesday July 9, 2024, 1:00 pm.

Problem 1 (Robustness of regularized least squares)**.** Suppose our goal is to recover a vector $\mathbf{x}$ from a noisy measurement $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}, m < n$, is a generic measurement matrix and $\mathbf{e}$ is additive noise. In this problem, we numerically study the behavior of the ℓ_2-regularized least-squares estimator

$$\hat{\mathbf{x}}_{2,\lambda}(\mathbf{y}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \frac{\lambda}{2}\|\mathbf{x}\|_2^2,$$

and the ℓ_1-regularized least-squares estimator

$$\hat{\mathbf{x}}_{1,\lambda}(\mathbf{y}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{x}\|_1,$$

under random (average-case) perturbations.

(a) Give a closed form expression of $\hat{\mathbf{x}}_{2,\lambda}$ in terms of the singular values and vectors of the matrix $\mathbf{A}$.

(b) Now we want to numerically investigate the reconstruction error of the estimator $\hat{\mathbf{x}}_{2,\lambda}$ under random perturbations. Towards this end, choose $\mathbf{A} \in \mathbb{R}^{500 \times 2000}$ as a Gaussian random matrix with $\mathcal{N}(0, 1/500)$ iid entries, and choose $\mathbf{x}$ as an 50-sparse vector with non-zeros drawn from the distribution $\mathcal{N}(0, 1/50)$. Plot the reconstruction mean-squared error $\|\hat{\mathbf{x}}_{2,\lambda}(\mathbf{A}\mathbf{x} + \mathbf{e}) - \mathbf{x}\|_2^2$ for a random noise vector $\mathbf{e}$ with iid $\mathcal{N}(0, 0.5/500)$ entries as a function of the regularization parameter λ.

(c) Repeat the steps from part (b) for the ℓ_1-regularized least-squares estimator $\hat{\mathbf{x}}_{1,\lambda}$. Note that there is no closed-form expression of the estimator $\hat{\mathbf{x}}_{1,\lambda}$ (why?), so you have to resort to optimization to approximate it. To this end, you can for example use the ISTA method as discussed in Problem 2 of Homework 2.

Problem 1 (Robustness of regularized least squares)**.** Suppose our goal is to recover a vector $\mathbf{x}$ from a noisy measurement $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}, m < n$, is a generic measurement matrix and $\mathbf{e}$ is additive noise. In this problem, we numerically study the behavior of the ℓ_2-regularized least-squares estimator

$$\hat{\mathbf{x}}_{2,\lambda}(\mathbf{y}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \frac{\lambda}{2}\|\mathbf{x}\|_2^2,$$

and the ℓ_1-regularized least-squares estimator

$$\hat{\mathbf{x}}_{1,\lambda}(\mathbf{y}) = \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda\|\mathbf{x}\|_1,$$

under random (average-case) perturbations.

(a) Give a closed form expression of $\hat{\mathbf{x}}_{2,\lambda}$ in terms of the singular values and vectors of the matrix $\mathbf{A}$.

$$L(x) = \frac{1}{2}\|Ax - y\|_2^2 + \frac{\lambda}{2}\|x\|_2^2$$

$$\nabla_x L(x) = A^T(Ax - y) + \lambda x \stackrel{!}{=} 0$$

$$(A^TA + I\lambda)x = A^Ty$$

$$x = (A^TA + \lambda I)^{-1} A^T y$$

Assume $A = U\Sigma V^T$

$$A^TA = U\Sigma^2 U^T$$

$$\boxed{x = (U\Sigma^2U^T + \lambda I)^{-1} \; V\Sigma U^T y}$$

assume $I = UU^T$: $\quad x = \left(U(\Sigma^2 + \lambda I)U^T\right)^{-1} V\Sigma U^T y$

b) See notebook

(c) Repeat the steps from part (b) for the ℓ_1-regularized least-squares estimator $\hat{\mathbf{x}}_{1,\lambda}$. Note that there is no closed-form expression of the estimator $\hat{\mathbf{x}}_{1,\lambda}$ (why?), so you have to resort to optimization to approximate it. To this end, you can for example use the ISTA method as discussed in Problem 2 of Homework 2.

It doesn't have a closed-form solution because $\|\cdot\|_1$ norm is not differentiable.

$$\Rightarrow g(x) = \lambda \|x\|_1 = \lambda \sum_{i=1}^{n} |x_i|$$

The absolute value is not differentiable at $x_i = 0$

$$\frac{d\,|x_i|}{dx_i} = \begin{cases} 1 & x_i > 0 \\ -1 & x_i < 0 \\ 0 & x_i = 0 \end{cases}$$

(See notebook)