MSc. Data Science & Artificial Intelligence

Inverse Problems in Image Processing

Faisal JAYOUSI, Laure BLANC-FERAUD & Luca CALATRONI

Assignment 1

Author: Joris LIMONIER

joris.limonier@gmail.com

Due: February 20, 2023

Contents

1 Exercise 1

Let f be given by:

$$f(x) = \frac{1}{2}\|Ax - y\|_2^2 \tag{1}$$

we want to compute the gradient of f. We have for a given direction $v \in \mathbb{R}^n$:

$$\begin{aligned}
\nabla_v f(x) &= \lim_{\varepsilon\to 0} \frac{f(x+\varepsilon v) - f(x)}{\varepsilon} \\
&= \lim_{\varepsilon\to 0} \frac{\|A(x+\varepsilon v) - y\|^2 - \|Ax - y\|^2}{2\varepsilon} \\
&= \lim_{\varepsilon\to 0} \frac{\|(Ax-y) + A\varepsilon v\|^2 - \|Ax - y\|^2}{2\varepsilon} \\
&= \lim_{\varepsilon\to 0} \frac{\|Ax-y\|^2 + 2\varepsilon < Ax - y, Av > + \varepsilon^2\|Av\|^2 - \|Ax - y\|^2}{2\varepsilon} \\
&= \lim_{\varepsilon\to 0} (Ax-y)^T Av + \underbrace{\frac{\varepsilon}{2}\|Av\|^2}_{\to 0} \\
&= \left\langle A^T(Ax-y), v \right\rangle,
\end{aligned}$$

We see that we obtain a scalar product with v on one side, as we wished. The other side of the scalar product (*i.e.* $A^T(Ax-y)$) corresponds to the gradient of f in the direction v. In other words, this is the change that will occur when we take a small step in the direction of v.

2 Exercise 2

2.1 Proof of the Lipschitz continuity of the gradient

Recall that a function g : is said to be L-Lipschitz continuous if $\forall\ x_1, x_2 \in \mathbb{R}^n$:

$$\|g(x_1) - g(x_2)\| \leq L\|x_1 - x_2\|. \tag{2}$$

In particular for $g \equiv \nabla f$, we have:

$$\begin{aligned}
\|\nabla f(x_1) - \nabla f(x_2)\| &= \|A^T(Ax_1 - y) - A^T(Ax_2 - y)\| \\
&= \|A^T A(x_1 - x_2)\| \\
&\leq \|A^T A\| \cdot \|x_1 - x_2\|
\end{aligned}$$

We therefore have that that ∇f is L-Lipschitz continuous, with $L := \|A^T A\|$.

2.2 Computation of the Lipschitz constant

Note that $\|A^T A\|$ is the matrix norm of $A^T A$, which is the largest singular value of $A^T A$. Let us call $\sigma_{max}(M)$ the largest singular value of a given matrix M. Recall that the singular value decomposition of M is given by:

$$M = U\Sigma V^T \tag{3}$$

where U and V are orthogonal matrices and Σ is a diagonal matrix with the singular values of M on the diagonal, which we assume to be sorted in decreasing order (without loss of generality, thanks to potential reordering of the rows, columns of U, V respectively). We have:

$$\begin{aligned}
\|A^T A\| &= \sigma_{max}(A^T A) \\
&= \sigma_{max}(U\Sigma \underbrace{V^T V}_{=Id} \Sigma U^T) \\
&= \sigma_{max}(U\Sigma^2 U^T) \\
&= \sigma_{max}(\Sigma^2) \\
&= \sigma_{max}(\Sigma)^2 \\
&= \sigma_{max}(U\Sigma V^T)^2 \\
&= \|A\|^2
\end{aligned}$$

3 Exercise 3

4 Exercise 4