Exercises
Computational Intelligence Lab
SS 2020

Machine Learning Institute
Dept. of Computer Science, ETH Zürich
Prof. Dr. Thomas Hofmann
Web http://da.inf.ethz.ch/cil

Series 11, May 22-23, 2020
(Dictionary Learning and Compressed Sensing)

Problem 1 (Sparse coding with overcomplete dictionary):

Given a signal $\mathbf{x} = (3, 1, -2) \in \mathbb{R}^3$ and an overcomplete dictionary $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4] \in \mathbb{R}^{3\times 4}$,

$$\mathbf{U} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix},$$

find the sparse representation $\mathbf{z}$ of the signal $\mathbf{x}$ with $\|\mathbf{z}\|_0 \le 2$.

a. Find the atom $\mathbf{u}^{(1)}$ that minimize the reconstruction error $\|\mathbf{x} - \hat{\mathbf{x}}^{(0)}\|_2^2$ where $\hat{\mathbf{x}}^{(0)} = z^{(1)}\mathbf{u}^{(1)}$, and compute the residual $\mathbf{r}^{(1)} = \mathbf{x} - \hat{\mathbf{x}}^{(0)}$.
Solution: The atom $\mathbf{u}^{(1)}$ that minimizes the reconstruction error $\|\mathbf{x} - z^{(1)}\mathbf{u}^{(1)}\|$ is the atom that is best correlated $\mathbf{x}$. The correlation between the signal and the atoms in the dictionary are as following,

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{u}_1 \rangle &= \frac{2}{\sqrt{3}} \\
\langle \mathbf{x}, \mathbf{u}_2 \rangle &= -\frac{4}{\sqrt{3}} \\
\langle \mathbf{x}, \mathbf{u}_3 \rangle &= 0 \\
\langle \mathbf{x}, \mathbf{u}_4 \rangle &= 2\sqrt{3}.
\end{aligned}$$

Since the absolute correlation coefficient between the atom $\mathbf{u}_4$ and the signal $\mathbf{x}$ has the largest value, $\hat{\mathbf{x}}^{(0)} = \langle \mathbf{x}, \mathbf{u}_4 \rangle \cdot \mathbf{u}_4 = 2\sqrt{3} \cdot \mathbf{u}_4$ minimizes $\|\mathbf{x} - \hat{\mathbf{x}}^{(0)}\|_2^2$. And the residual $\mathbf{r}^{(1)} = \mathbf{x} - \hat{\mathbf{x}}^{(0)} = (1, -1, 0)$

b. Find the atom $\mathbf{u}^{(2)}$ that minimize the reconstruction error $\|\mathbf{r}^{(1)} - \hat{\mathbf{x}}^{(1)}\|_2^2$ where $\hat{\mathbf{x}}^{(1)} = z^{(2)}\mathbf{u}^{(2)}$.
Solution: Similarly, we want to find the atom best correlated with $\mathbf{r}^{(1)}$ among the remaining atoms $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$. The correlation coefficients between the atoms and the residual are

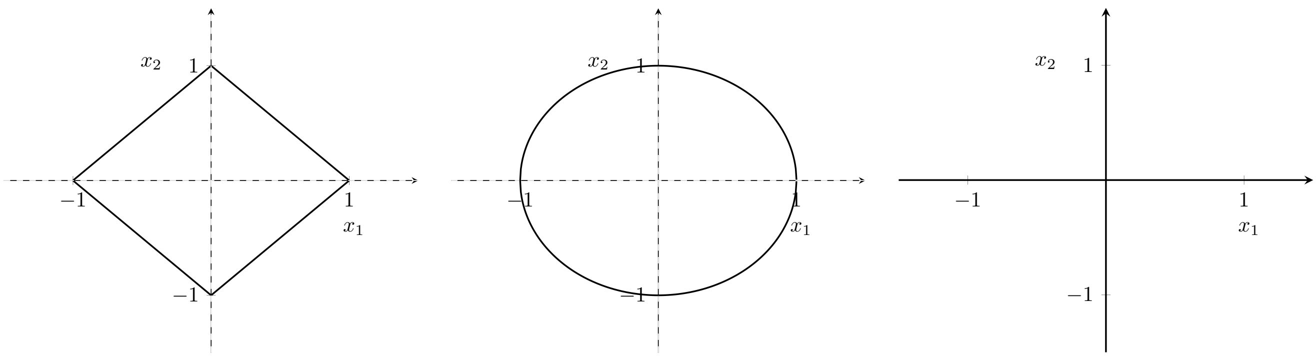
$$\begin{aligned}
\langle \mathbf{r}^{(1)}, \mathbf{u}_1 \rangle &= 0 \\
\langle \mathbf{r}^{(1)}, \mathbf{u}_2 \rangle &= -\frac{2}{\sqrt{3}} \\
\langle \mathbf{r}^{(1)}, \mathbf{u}_3 \rangle &= \frac{2}{\sqrt{3}}.
\end{aligned}$$

Because the aboslute correlation coefficient values of $\mathbf{u}_2$ and $\mathbf{u}_3$ are the same, either $\hat{\mathbf{x}}^{(1)} = -\frac{2}{\sqrt{3}}\mathbf{u}_2$ or $\hat{\mathbf{x}}^{(1)} = \frac{2}{\sqrt{3}}\mathbf{u}_3$ can minimize $\|\mathbf{r}^{(1)} - \hat{\mathbf{x}}^{(1)}\|_2^2$.

c. Write down the sparse representation $\mathbf{z}$ of signal $\mathbf{x}$.
Solution: The sparse representations $\mathbf{z}$ that satisfies $\|\mathbf{z}\|_0 \le 2$ are $(0, 0, 0, 2\sqrt{3})$, $(0, 0, \frac{2}{\sqrt{3}}, 2\sqrt{3})$ and $(0, -\frac{2}{\sqrt{3}}, 0, 2\sqrt{3})$.

Problem 2 (Compressed Sensing):

a. Map each of the three equations $\|\mathbf{x}\|_2 = 1$, $\|\mathbf{x}\|_1 = 1$, and $\|\mathbf{x}\|_0 = 1$ to a plot among a., b., or c. on the following figure. Note that $\mathbf{x}$ is s 2D vector with coordinates x_1 and x_2 (i.e. $\mathbf{x} = \left[x_1, x_2\right]$).



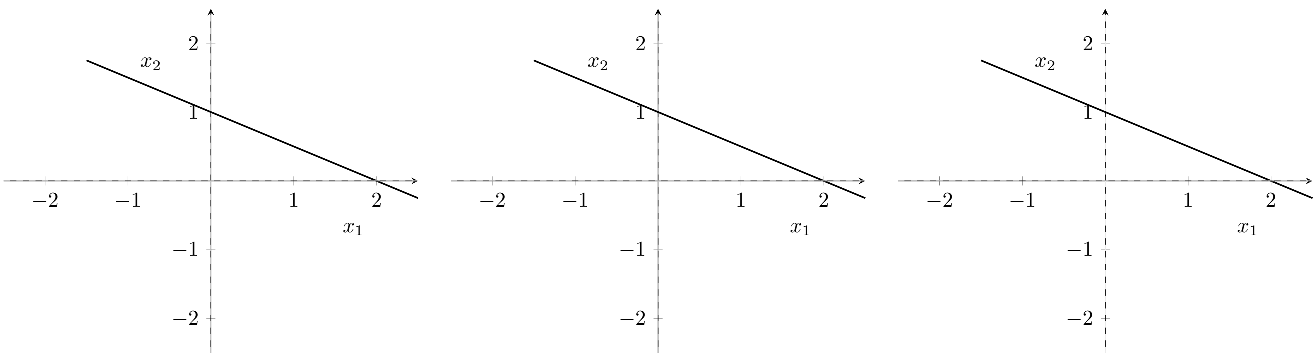
a. $\|\mathbf{x}\|_1$ b.$\|\mathbf{x}\|_2$ c. $\|\mathbf{x}\|_0$

b. Show the solution of each optimization problem on plots a., b., and c. of the following figure.

$\min \|\mathbf{x}\|_2$ Subject to $\frac{1}{2}x_1 + x_2 = 1$

$\min \|\mathbf{x}\|_1$ Subject to $\frac{1}{2}x_1 + x_2 = 1$

$\min \|\mathbf{x}\|_0$ Subject to $\frac{1}{2}x_1 + x_2 = 1$



a. b. c.

Solution:

a. The answer will be the closest point on the line to the origin, i.e.

$$\frac{1}{2}\mathbf{x}_1 + \mathbf{x}_2 = 1 \leftrightarrow \mathbf{x}_1 = 2 - 2\mathbf{x}_2 \tag{1}$$

$$\min\{\mathbf{x}_1^2 + \mathbf{x}_2^2\} = \min\{(2 - 2\mathbf{x}_2)^2 + \mathbf{x}_2^2\} \tag{2}$$

$$\frac{d}{d\mathbf{x}_2}\left[(2 - 2\mathbf{x}_2)^2 + \mathbf{x}_2^2\right] \stackrel{!}{=} 0 \leftrightarrow \mathbf{x}_2 = 0.8, \mathbf{x}_1 = 0.4 \tag{3}$$

b. $\mathbf{x}_1 = 0, \mathbf{x}_2 = 1$ c. two solutions $[\mathbf{x}_1 = 0, \mathbf{x}_2 = 1], [\mathbf{x}_1 = 2, \mathbf{x}_2 = 0]$

c. We can formulate the above three optimization problem as

$$\min \|\mathbf{x}\|_p$$
$$\text{subject to } \frac{1}{2}x_1 + x_2 = 1,$$

where $p \in \{0, 1, 2\}$. Mark the right sentence using your previous answers.

[ ] Solutions of the constrained problems have intersection for $p = 1$ and $p = 0$.

[ ] Solutions of the constrained problems have intersection for $p = 2$ and $p = 0$.

Solution:

Solutions of the constrained problems have intersection for $p = 1$ and $p = 0$.

Problem 3 (Compressed Sensing):

Please find the iPython notebook `Compressed_sensing.ipynb` from github.com/dalab/lecture_cil_public/tree/master/exercises/ex9

answer the question in this file.

Solution: Please find the solution in the same directory.

Problem 4 (Matching Pursuit Algorithm):

In the last tutorial session, you have seen that the matching pursuit algorithm converges. In this exercise, you will show some limitations of the algorithm.

a. Find an overcomplete dictionary and a vector $\mathbf{x}$ such that the approximation $\mathbf{x}$ resulting from the matching pursuit algorithm will never exactly equal $\mathbf{x}$ no matter the number of iterations. Note that this implies that at least one atom will be selected multiple times.

Hint: Look for $\mathbf{U} \in \mathbb{R}^{2\times 3}$ such that MP keeps alternating between $\mathbf{U}_{\bullet 1}$ and $\mathbf{U}_{\bullet 2}$

Solution: This exercise was phrased in a slightly confusing way: We want to find an example for which Matching Pursuit does not converge for a finite number of iterations. It will converge after ∞ steps as was proven in last week's tutorial.
Let's choose

$$\mathbf{U} = \begin{pmatrix} 1 & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{1}{2} \end{pmatrix}$$

and let us first consider a general signal of the form $\mathbf{x} = \begin{pmatrix} 0 \\ z \end{pmatrix}$ for $z \in \mathbb{R}$. In order to perform the first step of Matching Pursuit we need to calculate the inner products:

$$|\mathbf{U}_{\bullet 1}^T \mathbf{x}| = 0, \quad |\mathbf{U}_{\bullet 2}^T \mathbf{x}| = |z|\frac{\sqrt{2}}{2}, \quad |\mathbf{U}_{\bullet 3}^T \mathbf{x}| = \frac{|z|}{2}$$

We hence choose the second atom $\mathbf{U}_{\bullet 2}$ and obtain the following residual

$$\mathbf{r}^{(1)} = \mathbf{x} - \left(\mathbf{x}^T \mathbf{U}_{\bullet 2}\right) \mathbf{U}_{\bullet 2} = \begin{pmatrix} -\frac{z}{2} \\ \frac{z}{2} \end{pmatrix}$$

Now the second step of Matching Pursuit leads to the inner products

$$|\mathbf{U}_{\bullet 1}^T \mathbf{r}^{(1)}| = |y|, \quad |\mathbf{U}_{\bullet 2}^T \mathbf{r}^{(1)}| = 0, \quad |\mathbf{U}_{\bullet 3}^T \mathbf{r}^{(1)}| = \frac{\sqrt{3}-1}{4}|y|$$

We thus choose the first atom $\mathbf{U}_{\bullet 1}$ which leads to the following residual

$$\mathbf{r}^{(2)} = \mathbf{r}^{(1)} - \left(\mathbf{U}_{\bullet 1}^T \mathbf{r}^{(1)}\right) \mathbf{U}_{\bullet 1} = \begin{pmatrix} 0 \\ \frac{z}{2} \end{pmatrix}$$

Observe now that we are again in the same setting as we were at the start with $\mathbf{x} = \begin{pmatrix} 0 \\ z \end{pmatrix}$.
One can now show easily by induction that the residual at step $2n$ and $2n+1$ is given by

$$\mathbf{r}^{(2n)} = \begin{pmatrix} 0 \\ \frac{z}{2^n} \end{pmatrix}, \quad \mathbf{r}^{(2n+1)} = \begin{pmatrix} -\frac{z}{2^{n+1}} \\ \frac{z}{2^{n+1}} \end{pmatrix}$$

Thus, choosing e.g. $\mathbf{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, for any finite step n it holds that $\mathbf{r}^{(n)} \neq \mathbf{0}$. On the other hand, if $n \to \infty$, we have $\mathbf{r}^{(\infty)} = \mathbf{0}$ as expected.

b. Find an instance where the sparse representation returned by matching pursuit (assuming that after some number of iterations, the approximation is perfect) is not optimal, i.e. there is a different representation for which the 0-norm is strictly smaller.

Hint: Look for $\mathbf{U} \in \mathbb{R}^{2\times 3}$ and $\mathbf{x} \in \mathbb{R}^2$ such that you can find a representation $\mathbf{z}$ "by hand" with $||\mathbf{z}||_0 = 2$ but MP returns a representation $\mathbf{z}_{\text{MP}}$ with $||\mathbf{z}_{\text{MP}}||_0 = 3$

Solution: Let's choose

$$\mathbf{U} = \begin{pmatrix} 1 & 0 & \frac{\sqrt{2}}{2} \\ 0 & 1 & \frac{\sqrt{2}}{2} \end{pmatrix} \text{ and } \mathbf{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

We see that we can come up "by hand" with the representation

$$\mathbf{x} = 2 \cdot \mathbf{U}_{\bullet 1} + 1 \cdot \mathbf{U}_{\bullet 2} + 0 \cdot \mathbf{U}_{\bullet 3} \implies \mathbf{z} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

Let us now calculate what Matching Pursuit returns. The inner products are given by

$$|\mathbf{U}_{\bullet 1}^T \mathbf{x}| = 2, \quad |\mathbf{U}_{\bullet 2}^T \mathbf{x}| = 1, \quad |\mathbf{U}_{\bullet 3}^T \mathbf{x}| = \frac{3}{\sqrt{2}}$$

We hence choose the third atom $\mathbf{U}_{\bullet 3}$ and we obtain the residual

$$\mathbf{r}^{(1)} = \mathbf{x} - \left(\mathbf{x}^T \mathbf{U}_{\bullet 1}\right) \mathbf{U}_{\bullet 1} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$$

The next step gives

$$|\mathbf{U}_{\bullet 1}^T \mathbf{r}^{(1)}| = \frac{1}{2}, \quad |\mathbf{U}_{\bullet 2}^T \mathbf{r}^{(1)}| = \frac{1}{2}, \quad |\mathbf{U}_{\bullet 3}^T \mathbf{r}^{(1)}| = 0$$

We are hence free to choose between the first atom $\mathbf{U}_{\bullet 1}$ and the second $\mathbf{U}_{\bullet 2}$. Let's go with the first one which results in the residual

$$\mathbf{r}^{(2)} = \mathbf{r}^{(1)} - \left(\mathbf{U}_{\bullet 3}^T \mathbf{r}^{(1)}\right) \mathbf{U}_{\bullet 3} = \begin{pmatrix} 0 \\ -\frac{1}{2} \end{pmatrix}$$

The third step now leads to

$$|\mathbf{U}_{\bullet 1}^T \mathbf{r}^{(2)}| = 0, \quad |\mathbf{U}_{\bullet 2}^T \mathbf{r}^{(2)}| = \frac{1}{2}, \quad |\mathbf{U}_{\bullet 3}^T \mathbf{r}^{(2)}| = \frac{\sqrt{2}}{4}$$

Thus we go with the second atom $\mathbf{U}_{\bullet 2}$ which leads to the residual

$$\mathbf{r}^{(3)} = \mathbf{r}^{(2)} - \left(\mathbf{U}_{\bullet 3}^T \mathbf{r}^{(2)}\right) \mathbf{U}_{\bullet 3} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Matching pursuit hence converges after three steps and the algorithm terminates. We obtain the representation

$$\mathbf{z}_{MP} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{3}{\sqrt{2}} \end{pmatrix}$$

Notice that $||\mathbf{z}_{MP}||_0 = 3 > ||\mathbf{z}||_0 = 2$ and as a result, Matching Pursuit did not find the sparsest representation.