ELE 535 - Machine Learning and Pattern Recognition

Fall 2018

HOMEWORK 7: Theory

Q1 **Sparse Representation in an ON Basis.** Let $r \leq n$ and $Q \in \mathbb{R}^{n \times r}$ have orthonormal columns.
(a) Find a solution of the following sparse approximation problem and determine if the solution is unique.

$$\begin{aligned} \min_{w \in \mathbb{R}^w} \quad & \|y - Qw\|_2^2 \\ \text{subject to:} \quad & \|w\|_0 \leq k. \end{aligned}$$

(b) Now let the columns of $X \in \mathbb{R}^{n \times m}$ be a centered set of unlabelled training data and the columns of $Q \in \mathbb{R}^{n \times r}$ be the left singular vectors of a compact SVD of X. In this context, interpret the solution of the above problem.

Q2 Let

$$M = \begin{bmatrix} \mathbf{e}_1 & \frac{1}{\sqrt{2}}(\mathbf{e}1 + \mathbf{e}_2) & \mathbf{e}_3 & \frac{1}{\sqrt{3}}(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) \end{bmatrix},$$

where $\mathbf{e}_i$ denotes the i-th standard basis vector in $\mathbb{R}^n$.

(a) Show that the columns of M are linearly dependent.
(b) Determine $\text{spark}(M)$.
(c) Determine the mutual coherence $\mu(M)$.

Q3 Let $A \in \mathbb{R}^{m \times n}$ with $\text{rank}(A) = m < n$, and $y \in \mathbb{R}^m$. We seek the sparsest solution of $Ax = y$:

$$\min_{x \in \mathbb{R}^n} \|x\|_0, \quad \text{s.t.} \quad Ax = y. \tag{1}$$

The convex relaxation of this problem is called Basis Pursuit:

$$\min_{x \in \mathbb{R}^n} \|x\|_1, \quad \text{s.t.} \quad Ax = y. \tag{2}$$

Show that (2) is equivalent to the linear program:

$$\begin{array}{lrcl} \min_{x,z \in \mathbb{R}^n} & \mathbf{1}^T z & & \\ \text{s.t.} & Ax & = & y \\ & x - z & \leq & \mathbf{0} \\ & -x - z & \leq & \mathbf{0}. \end{array}$$

Q4 One way to create a dictionary is to combine known ON bases. Here we explore combining the standard basis with the Haar wavelet basis.
The Haar wavelet basis consists of $n = 2^p$ ON vectors in $\mathbb{R}^n$. These can be arranged into the

columns of an orthogonal matrix H_p with

$$H_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \qquad H_2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$H_3 = \begin{bmatrix} \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{4}} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{4}} & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{4}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{8}} & \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{4}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{4}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 0 & \frac{1}{\sqrt{4}} & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 0 & -\frac{1}{\sqrt{4}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{8}} & -\frac{1}{\sqrt{8}} & 0 & -\frac{1}{\sqrt{4}} & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

The columns are arranged in groups. The first group consists of the vector $\frac{1}{\sqrt{n}}\mathbf{1}_n$, and second consists of the vector taking the value $\frac{1}{\sqrt{n}}$ in the first half, and $-\frac{1}{\sqrt{n}}$ in the second half. Subsequent groups of vectors are derived by subsampling by 2, scaling by $\sqrt{2}$, and translating. This is illustrated above for $p = 1, 2, 3$.

Form a dictionary $D \in \mathbb{R}^{n \times 2n}$ by setting $D = [I_n, H_p]$ with $n = 2^p$. The matrix H_p is the Haar matrix of size $n = 2^p$.

Show that

(a) For $p = 1$, $\mathrm{spark}(D) = 3$ and $\mu(D) = 1/\sqrt{2}$.
(b) For all $p > 1$, determine $\mathrm{spark}(D)$ and $\mu(D)$.
(c) For a given $y \in \mathbb{R}^n$, we seek the sparest solution of $y = Dw$. What condition on y is sufficient to ensure the sparsest solution is unique?

ONLY FOUR QUESTIONS THIS WEEK.