

DEEP LEARNING AND INVERSE PROBLEMS  
SUMMER 2025  
Reinhard Heckel

**Chapter 4 exercises**

Issued: Thursday May 6, 2025

Due: Thursday May 13, 2025.

**Problem 1** (Signal separation). Below is a question and the corresponding response from ChatGPT (generated with GPT4 in May 2023). Is the algorithm proposed by ChatGPT correct? If yes, can you help to answer how large the sparsity levels  $s_1$  and  $s_2$  can be for the algorithm to succeed? Provide references to results from the literature (e.g., from the book) in your answers.

Model: GPT-4

R

Suppose I have a signal  $x$  that is a superposition of a signal that is  $s_1$ -sparse in a  $n \times n$  Gaussian random matrix  $A_1$ , and a signal that is  $s_2$ -sparse in another, independently chosen  $n \times n$  Gaussian random matrix  $A_2$ . How can I separate the two parts with basis pursuit, and how large can  $s_1$  and  $s_2$  be?



Basis pursuit is an optimization technique used to solve linear inverse problems with sparse solutions. It aims to find the sparsest solution by minimizing the L1-norm of the solution vector. In this problem, the signal  $x$  is a superposition of two sparse signals,  $x_1$  and  $x_2$ , in two different bases  $A_1$  and  $A_2$ .



To separate the two parts using basis pursuit, you can formulate the problem as follows:

minimize  $\|x_1\|_1 + \|x_2\|_1$   
subject to  $x = A_1 x_1 + A_2 x_2$

Here,  $\|\cdot\|_1$  denotes the L1-norm of the vector.

You can solve this optimization problem using algorithms like LASSO, Basis Pursuit Denoising (BPDN), or other convex optimization techniques.

Now, regarding the sparsity levels  $s_1$  and  $s_2$ , the success of the recovery mainly depends on the Restricted Isometry Property (RIP) and mutual coherence between the bases  $A_1$  and  $A_2$ . In general, the smaller the mutual coherence and the better the RIP, the higher the sparsity levels you can achieve.

However, a precise bound on  $s_1$  and  $s_2$  is difficult to provide without more information about the matrices  $A_1$  and  $A_2$ . In practice, you may need to use tools like coherence analysis or the RIP to analyze your specific case and determine the appropriate sparsity levels. Keep in mind that recovering the original signals becomes more challenging as the sparsity levels increase.

**Problem 2** (Compressive sensing with spikes and random vectors). In this problem we consider recovery of a signal that is sparse in the concatenation of the identity basis and a random basis, i.e., in the dictionary

$$\mathbf{D} = [\mathbf{B}, \mathbf{I}] \in \mathbb{R}^{n \times 2n}.$$

Here,  $\mathbf{B} \in \mathbb{R}^{n \times n}$  is random Gaussian matrix, with entries iid  $\mathcal{N}(0, 1/n)$ , and  $\mathbf{I} \in \mathbb{R}^{n \times n}$  contains

the standard basis vectors, i.e., the  $k$ -th column of  $\mathbf{I}$  is the unit vector  $\mathbf{e}_k$  with  $k$ -th entry equal to one.

1. First, we are given the measurement

$$\mathbf{y} = [\mathbf{b}_0, \mathbf{e}_0, \dots, \mathbf{e}_{n-1}] \mathbf{x},$$

where  $\mathbf{x}$  is  $s$ -sparse. Furthermore, suppose the support set of  $\mathbf{x}$ , i.e., the position of the non-zeros is known. How large can  $s$  be such that recovery of every  $s$ -sparse vector from the corresponding measurement  $\mathbf{y}$  is possible, with probability one, and why?

Next, we study recovery of  $\mathbf{x}$  from the measurement

$$\mathbf{y} = \mathbf{D}\mathbf{x}.$$

Towards this goal, we first study the incoherence of the matrix  $\mathbf{D}$ . Recall from class, that a matrix  $\mathbf{A}$  with unit norm columns is  $\mu$ -incoherent if  $\mu$  is the smallest number such that for all pairs of columns  $\mathbf{a}_i, \mathbf{a}_j$ ,  $i \neq j$  of  $\mathbf{A}$ ,

$$|\langle \mathbf{a}_i, \mathbf{a}_j \rangle| \leq \mu.$$

Also recall that for a Gaussian random variable  $x \sim \mathcal{N}(0, 1)$ , we have that, for  $\beta \geq \frac{1}{2\pi}$  that

$$\mathbb{P}[x \geq \beta] \leq e^{-\beta^2/2}.$$

2. What is the distribution of  $\langle \mathbf{e}_i, \mathbf{b}_j \rangle$ , and given this result, what is bound on the probability that the absolute values of this random variable exceeds a constant  $\beta'$ ?
3. Building on the result from the previous part, what is the coherence parameter of the matrix  $\mathbf{D}$ , with high probability (say with probability at least  $1 - \delta$ , where  $\delta > 0$  is a parameter)? Since  $\|\mathbf{b}_i\|_2 \approx 1$ , with high probability, you can assume  $\|\mathbf{b}_i\|_2 = 1$  for simplicity.
4. Recall from class that provided a matrix  $\mathbf{A}$  is  $\mu$ -incoherent, and  $\mathbf{x}$  is  $s$ -sparse with  $s < \frac{1}{2\mu}$ , then  $\ell_1$ -minimization provably recovers the vector  $\mathbf{x}$  from the measurements  $\mathbf{y} = \mathbf{A}\mathbf{x}$ . Based on this result, what is the maximum sparsity under which recovery with  $\ell_1$ -minimization provably succeeds? Is this result pessimistic, i.e., is there a matrix that allows the sparsity  $s$  to be significantly larger, if yes, specify that matrix and specify what the allowed level of sparsity is.