

Declaración de Originalidad

Ricardo Beneyto Navarro, autor del Trabajo de Fin de Máster “Compressed Sensing y la reconstrucción de señales”, bajo la tutela de los profesores Gustavo Garrigós y Matías Raja, declara que el trabajo que presenta es original, en el sentido de que ha puesto el mayor empeño en citar debidamente todas las fuentes utilizadas, y que la obra no infringe el copyright de ninguna persona.

En Murcia, a 5 de junio de 2022.

Fdo: Ricardo Beneyto Navarro

Índice general

| | Página |
|--|-----------|
| 1 Introducción | 1 |
| 1 Señales sparse | 1 |
| 2 Resumen del trabajo | 3 |
| 2 Reconstrucción de señales sparse | 9 |
| 1 Minimización ℓ_0 | 9 |
| 2 Método de los Coeficientes de Fourier | 12 |
| 3 Normas ℓ_p | 17 |
| 3 Compressed Sensing mediante la Minimización ℓ_1 | 21 |
| 1 La Propiedad de Isometría Restringida | 21 |
| 2 La Propiedad de Concentración de Medida | 26 |
| 3 Número de mediciones | 31 |
| 4 Distribución Subgaussiana | 35 |
| 1 Definición y momentos | 35 |
| 2 Concentración de Medida | 39 |
| 3 Caracterización y recíproco | 46 |
| 4 Lema de Johnson-Lindenstrauss | 49 |
| 5 Ejemplos | 50 |
| 5 Minimización ℓ_1 en Octave | 53 |
| 1 Ejemplo básico: Estabilidad y robustez | 54 |
| 2 Reconstrucción de imágenes. | 56 |
| Apéndice | 59 |
| A Matrices de Vandermonde | 59 |
| B Transformada de Fourier y Convolución Discretas | 60 |
| C Conceptos Básicos de Probabilidad | 64 |
| D Códigos de Octave | 67 |
| Referencias | 74 |

Índice de figuras

| | Página |
|--|---------------|
| 1.1 Ilustración del nivel de <i>sparse</i> de una señal real. | 2 |
| 1.2 Algunos de coeficientes de la base de Wavelets de la imagen (5.2.10). | 3 |
| 2.1 Reconstrucción exacta partir de $2s$ coeficientes de Fourier. | 15 |
| 2.2 Problemas de robustez en la reconstrucción de Fourier. | 16 |
| 2.3 Módulo de los primeros $2s$ coeficientes de la TFD de la señal. | 16 |
| 2.4 Problemas de estabilidad en la reconstrucción de Fourier (Pico en 180). | 17 |
| 2.5 Problemas de estabilidad en la reconstrucción de Fourier (Pico en 200). | 17 |
| 2.6 Ejemplo de los algoritmos de minimización con las normas ℓ_1 y ℓ_2 | 18 |
| 2.7 Regiones donde los algoritmos reconstruyen la señal de manera exacta. | 19 |
| 5.1 Ejemplo Básico: Señal original. | 54 |
| 5.2 Ejemplo Básico: Mejor y peor reconstrucción con $m = 80$ | 55 |
| 5.3 Ejemplo Básico: Robustez ante errores de medición. | 56 |
| 5.4 Ejemplo Básico: Estabilidad antes señales no <i>sparse</i> del todo. | 56 |
| 5.5 Fotografía de la superficie de la Luna. (5.1.09) | 57 |
| 5.6 Los coeficientes de la base de Wavelets de la imagen. | 57 |
| 5.7 Aproximación <i>s-sparse</i> de la imagen. | 58 |
| 5.8 Reconstrucción mediante Minimización ℓ_1 de la imagen. | 58 |

Capítulo 1

Introducción

Numerosos problemas prácticos consisten en medir información para poder analizarla y obtener, a partir de ella, algún resultado o conclusión. La mayoría de las veces, esta información viene dada en forma de una cadena finita de números reales, que de forma genérica se denomina “señal” y que representaremos mediante un vector x de \mathbb{R}^N .

Habrán situaciones, ya sea por el gran tamaño de la señal o porque el proceso de medición sea muy costoso o difícil, en las que no seremos capaces de medir la señal en su totalidad y solo dispondremos de cierta información. Nuestro objetivo será entonces tratar de reconstruir la señal completa a partir de ese número limitado de mediciones. En concreto, nos centraremos en el caso en que el proceso de medición sea lineal ya que así podremos reducir la reconstrucción a resolver un sistema de ecuaciones lineales.

En términos matemáticos tendremos $y = \Phi x$, donde $y \in \mathbb{R}^m$ es la información medida, $x \in \mathbb{R}^N$ es la señal original que queremos recuperar y Φ es una matriz real de tamaño $m \times N$ que modeliza el proceso de recabar información. Así, para recuperar x lo que haremos será intentar resolver este sistema lineal.

El álgebra lineal clásico nos dice que si $m < N$ entonces el sistema será indeterminado y existirán infinitas soluciones, dadas por el subespacio afín $x + \ker \Phi$. Es decir, sin ninguna información adicional es imposible recuperar unívocamente x a partir de y si $m < N$. Por ello, uno podría pensar que para poder reconstruir la señal haría falta que el número de mediciones, m , sea al menos tan grande como la longitud de la señal, N . Pero, como ya hemos dicho, esto no siempre ocurre por lo que será necesario estudiar soluciones alternativas.

En este trabajo, vamos a estudiar la teoría del *Compressed Sensing*, o Muestreo comprimido de señales, y veremos que bajo ciertas condiciones podremos recuperar las señales incluso con $m < N$. La clave estará en suponer que las señales con las que trabajamos cumplen la propiedad de ser *sparse*, es decir, la mayoría de sus componentes son ceros. En la siguiente sección daremos la definición rigurosa del concepto *sparse* y mostraremos ejemplos prácticos en los que las señales cumplen de forma natural esta propiedad.

1. Señales sparse

Estamos representando las señales como vectores en \mathbb{R}^N por lo que las podemos expresar en términos de una base ortonormal prefijada. Así, si $\{b_i\}_{i=1}^N$ es nuestra base, tendremos

$$x = \sum_{i=1}^N x_i b_i, \text{ donde } x_i = \langle x, b_i \rangle.$$

Con esto, y llamando $[N] = \{1, \dots, N\}$, definimos el soporte de la señal x , con respecto a la base $\{b_i\}$, como

$$\text{Sop } x := \{i \in [N] : x_i \neq 0\}.$$

Definición 1.1 (Sparse).

Diremos que una señal $x \in \mathbb{R}^N$ es s -sparse si es una combinación lineal de a lo sumo s vectores de la base, es decir, si $\text{card}(\text{Sop } x) \leq s$. Denotaremos por Σ_s el conjunto de todos los vectores s -sparse de \mathbb{R}^N .

El término *sparse* podría traducirse al castellano como escaso o disperso. No obstante, ninguna de estas palabras parece adecuarse mucho al significado que queremos transmitir por lo que durante el trabajo utilizaremos la palabra *sparse*.

En la práctica puede ocurrir que las señales que nos interesan no sean exactamente *sparse*, pero sí aproximadamente *sparse*, en el sentido de que, expresadas en una base adecuada, tendrán una gran cantidad de componentes muy pequeñas y, por tanto, se podrán aproximar muy bien por señales *sparse*. Bastaría hacer cero sus componentes más pequeñas y dejar el resto como en la original.

Para ilustrarlo tomaremos la fotografía 5.2.10 de la base de datos USC-SIPI Image Database [16]. Se trata de una imagen en escala de grises con resolución 512×512 píxeles y profundidad 8 bits. Es decir, para representarla tendremos un valor numérico entre 0 y 255 por cada píxel que nos indica su tono de gris, siendo 0 el negro y 255 el blanco.

Si vectorizamos la matriz de píxeles, la señal que nos queda tendrá tamaño $N = 512^2 = 262144$. Pero si la expresamos en términos de una base de Wavelets se puede ver que la mayoría de sus coeficientes son relativamente pequeños. Podemos incluso quedarnos solo con los 50000 coeficientes más grandes (menos del 20% de los datos), poner el resto a 0 y, aún así, la imagen reconstruida se asemejará a la original.

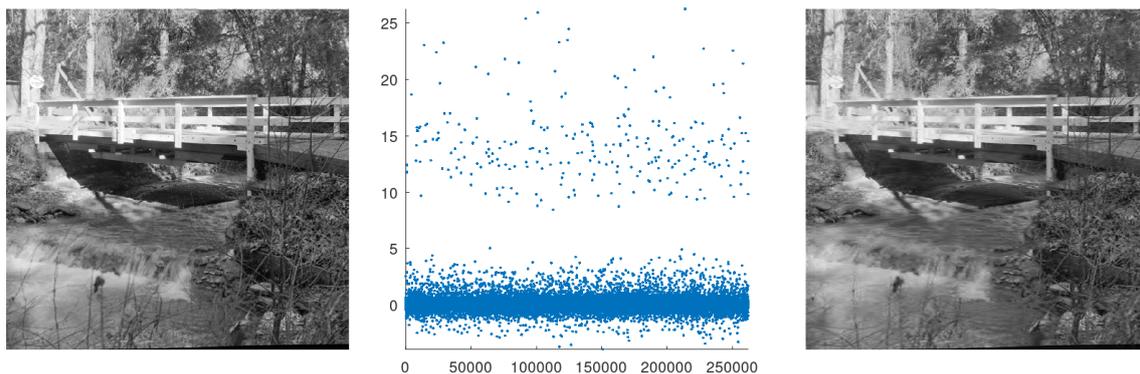


Figura 1.1: Izquierda: La imagen original (5.2.10). Centro: Los coeficientes de la base de Wavelets (permutados aleatoriamente para mejorar su visibilidad). Derecha: La imagen reconstruida a partir del vector con los 50000 coeficientes más grandes y el resto puestos a 0.

Puede parecer que los coeficientes están distribuidos por igual en la banda entorno al valor 0. No obstante, si hacemos zoom a la gráfica veremos que están mucho más concentrados que están en los valores pequeños. De hecho, en el zoom se muestra la franja de coeficientes que estamos poniendo a 0 en la aproximación *sparse* de la imagen.

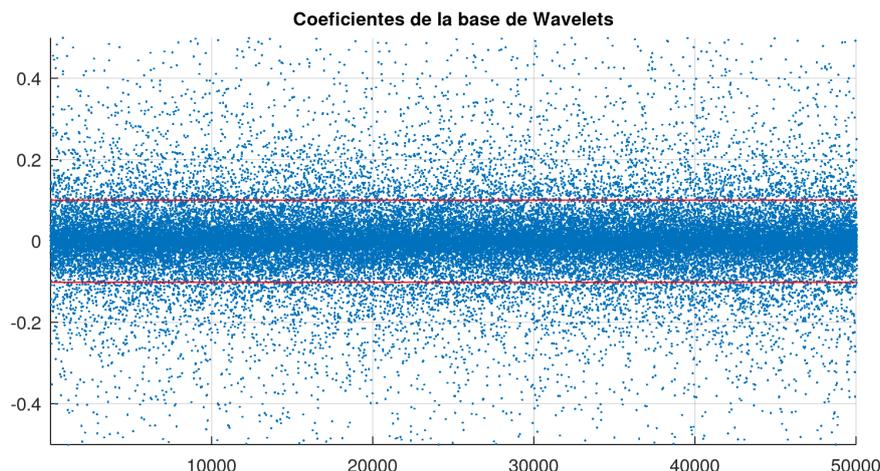


Figura 1.2: Algunos de coeficientes de la base de Wavelets de la imagen del puente (5.2.10). Los puntos entre las líneas rojas pasarán a valer 0 en la aproximación *sparse*.

Este ejemplo ilustra que las señales a menudo se pueden “comprimir” utilizando un número de parámetros mucho menor que su tamaño. Teniendo esto en cuenta, ya no parece tan extraño que seamos capaces de reconstruir las señales con un número de mediciones menor al tamaño, o sea, $m < N$.

Esta es una de las razones por las cuales el *Compressed Sensing* es una alternativa atractiva en la compresión de señales. El enfoque tradicional consistía en medir la señal al completo y luego comprimirla. Pero si ya sabemos que la señal es *sparse* y la mayoría de coeficientes se descartarán en la compresión, parece una pérdida de recursos tratar de medirla entera. Sería más razonable buscar un método que permita medir directamente una versión comprimida de la señal con un número menor de mediciones y luego tratar de reconstruir la señal original, en caso de necesitarla. De ahí que se le llame a la técnica *Compressed Sensing* o Muestreo comprimido de Señales.

2. Resumen del trabajo

Antes de terminar la introducción explicaremos en que consistirá el trabajo y como está distribuido.

Para empezar, en el Capítulo 2, desarrollaremos la técnica del Compressed Sensing y estudiaremos algunos métodos de reconstrucción. Primero nos centraremos en la Minimización l_0 , que consiste en buscar el vector más *sparse* posible pero que sea acorde a los datos. Es decir, si denotamos $\|x\|_0 := \text{card}(\text{Sop } x)$ trataremos de reconstruir la señal con las soluciones al problema de optimización

$$(P_0) : \left[\arg \min_{v \in \mathbb{R}^N} \|v\|_0 \text{ sujeto a } \Phi v = y \right],$$

donde $\arg \min_{v \in S} \|v\|_0 := \{v \in S : \|v\|_0 \leq \|u\|_0, \forall u \in S\}$, para $S \subseteq \mathbb{R}^N$.

Siguiendo los apuntes del Capítulo 2 del libro de S. Foucart y H. Rauhut [11] obtendremos el siguiente resultado.

Teorema (Ver 2.2). *Para cualquier entero $N \geq 2s$, existe una matriz Φ de tamaño $2s \times N$ tal que cualquier vector $x \in \Sigma_s$ se puede recuperar a partir de su vector de mediciones $y = \Phi x \in \mathbb{R}^{2s}$ como solución del problema (P_0) .*

Lo que quiere decir que con tan sólo $2s$ mediciones el método teórico de la Minimización ℓ_0 será capaz de reconstruir cualquier señal s -sparse de manera exacta. Este número de mediciones coincide con lo que nos sugiere la intuición, que para recuperar una señal s -sparse nos hará falta conocer tanto el valor de sus s componentes no nulas como las s posiciones de estas componentes.

No obstante, observaremos que tratar de resolver el problema (P_0) en general no es viable debido a los elevados tiempos de computación. Por suerte, en [11] también se propone un método de reconstrucción alternativo a partir de los primeros $2s$ coeficientes de la transformada de Fourier de la señal. Este método solucionará el problema del tiempo de computación pero, como veremos en un ejemplo, ni será robusto frente a errores de medición, ni estable frente a problemas con el nivel de *sparse* de la señal.

Tras ver que la Minimización ℓ_0 no funciona tan bien como esperábamos, buscaremos entre las normas ℓ_p algún método que posea algoritmos de resolución eficientes y que, además, sea robusto y estable. En concreto, compararemos las dos normas más conocidas, ℓ_1 y ℓ_2 . Veremos que, en general, la norma ℓ_2 no reconstruye señales *sparse*, y que la norma ℓ_1 sí que parece funcionar mucho mejor en ese sentido.

De esta manera, en el Capítulo 3, pasaremos a analizar el método de la Minimización ℓ_1 que, como antes, consistirá en reconstruir x a partir de las soluciones del problema

$$(P) : \left[\arg \min_{v \in \mathbb{R}^N} \|v\|_1 \text{ sujeto a } \Phi v = y \right].$$

Nuestro primer objetivo será probar que cualquier solución x^* del problema (P) recuperará de manera exacta la señal x si esta es suficientemente *sparse*. Para ello hay que pedir a la matriz Φ que cumpla la Propiedad de Isometría Restringida (RIP).

Definición. Para cada entero positivo s se define la constante de isometría de una matriz Φ de tamaño $m \times N$, como el menor número $\delta_s \geq 0$ tal que

$$(1 - \delta_s) \|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_s) \|x\|_2^2, \quad \forall x \in \Sigma_s. \quad (1.1)$$

Se dice que la matriz satisface la Propiedad de Isometría Restringida (RIP) de orden s si

$$0 \leq \delta_s < 1.$$

De hecho, siguiendo el artículo de E. J. Candès [7], seremos capaces de probar un resultado más general con el que veremos que el método ℓ_1 es de hecho estable y robusto si Φ cumple RIP. Para ello, modificaremos el problema (P) para que incluya ruido en las mediciones, es decir, consideramos el problema

$$(P^*) : \left[\arg \min_{v \in \mathbb{R}^N} \|v\|_1 \text{ sujeto a } \|y - \Phi v\|_2 \leq \varepsilon \right],$$

donde ε es una cota superior del ruido. Además, denotaremos por x_s la mejor aproximación s -sparse de x , que será el vector con los s coeficientes más grandes de x y el resto 0.

Teorema (Ver 3.2). *Si suponemos que $\delta_{2s} < \sqrt{2} - 1$ y $\|y - \Phi x\|_2 \leq \varepsilon$, entonces si x^* es solución del problema (P*) se cumple*

$$\|x^* - x\|_2 \leq \frac{c_0}{\sqrt{s}} \|x - x_s\|_1 + c_1 \varepsilon$$

para constantes $c_0, c_1 > 0$, que dependen de δ_{2s} .

Así, si suponemos que no cometemos errores de medición y que la señal es s -sparse tendremos $\varepsilon = 0$ y $x = x_s$ con lo que la reconstrucción será exacta.

Por supuesto, lo que nos interesa ahora es construir matrices de medición que cumplan la propiedad RIP. Basándonos en el artículo [2], demostraremos que las matrices aleatorias generadas a partir de una distribución que cumple la llamada propiedad de Concentración de Medida cumplirán también RIP con una probabilidad alta.

Definición. Decimos que una variable aleatoria, X , cumple la propiedad de Concentración de Medida si, para las matrices aleatorias $\Phi(\omega)$ de tamaño $m \times N$ cuyas entradas son realizaciones independientes de la variable X , se cumple: $\mathbb{E} [\|\Phi(\omega)x\|_2^2] = \|x\|_2^2$ para cualquier $x \in \mathbb{R}^N$ y para cualquier $0 < \varepsilon < 1$ existe una constante $c_0(\varepsilon) > 0$ tal que

$$\mathbb{P} \left(\left| \|\Phi(\omega)x\|_2^2 - \|x\|_2^2 \right| \geq \varepsilon \|x\|_2^2 \right) \leq 2e^{-m c_0(\varepsilon)}.$$

El resultado que probaremos es más general pero basta tomar $\delta < \sqrt{2} - 1$ para tener RIP.

Teorema (Ver 3.6). *Sea Φ una matriz aleatoria de tamaño $m \times N$ obtenida de acuerdo a una distribución que cumple la Concentración de Medida. Entonces, dado $0 < \delta < 1$, existen constantes $c_1, c_2 > 0$, que dependen solo de δ , tales que para cualquier $s \leq \frac{c_1 m}{\log(N/s)}$ con probabilidad $\geq 1 - 2e^{-c_2 m}$, $\Phi(\omega)$ cumple*

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2, \quad \forall x \in \Sigma_s.$$

Con la relación $s \leq \frac{c_1 m}{\log(N/s)}$ obtenemos el número de mediciones necesarias para recuperar las señales s -sparse con matrices de medición aleatorias. De hecho, con el Teorema 10.11 del libro [11] veremos que es óptimo independientemente de la matriz de medición escogida.

Teorema (Ver 3.8). *Dada una matriz $A \in \mathbb{R}^{m \times N}$, si todo vector $2s$ -sparse $x \in \mathbb{R}^N$ minimiza $\|z\|_1$ sujeto a que $Az = Ax$, entonces*

$$m \geq c_1 s \log \left(\frac{N}{c_2 s} \right), \quad \text{donde } c_1 = \frac{1}{\log(9)} \text{ y } c_2 = 4.$$

Cambiando $2s$ por s nos dará la cota para el número mínimo de mediciones necesarias para que el método de la Minimización ℓ_1 sea exacto. Decíamos que la cota del teorema anterior era óptima porque ambas tienen el mismo orden, $s \log(N/s)$.

Durante el Capítulo 4 nos centraremos en las variables aleatorias Subgaussianas utilizando, principalmente, los resultados del Capítulo 7 del libro [3].

Definición. Una variable aleatoria X diremos que es Subgaussiana si existe una constante $c > 0$ tal que

$$\mathbb{E}[e^{tX}] \leq e^{\frac{c^2 t^2}{2}}, \quad \forall t \in \mathbb{R}.$$

La denotaremos $X \sim \text{Sub}(c^2)$.

Lo primero que veremos es que, si $X \sim \text{Sub}(c^2)$, sus momentos de primer y segundo orden cumplen $\mathbb{E}[X] = 0$ y $\mathbb{E}[X^2] \leq c^2$. De hecho, definiremos una clase de variables aleatorias algo más restrictiva en las que se da la igualdad en la segunda expresión.

Definición. Una variable aleatoria X diremos que es Estrictamente Subgaussiana si existe una constante $\sigma > 0$ tal que $X \sim \text{Sub}(\sigma^2)$ y $\mathbb{E}[X^2] = \sigma^2$. La denotaremos $X \sim \text{SSub}(\sigma^2)$.

El objetivo principal de este capítulo es probar que las matrices aleatorias generadas a partir de variables Estrictamente Subgaussianas cumplirán la propiedad de Concentración de Medida y, por tanto, podrán utilizarse en la reconstrucción de señales.

Teorema (Ver 4.15). *Sea Φ una matriz $m \times N$ cuyas entradas $\Phi_{i,j}$ son independientes e idénticamente distribuidas con $\Phi_{i,j} \sim \text{SSub}(1/m)$. Entonces, para cualquier $\varepsilon \in (0, 1)$ y cualquier $x \in \mathbb{R}^N$, tendremos $\mathbb{E}[\|\Phi x\|_2^2] = \|x\|_2^2$ y*

$$\mathbb{P}(|\|\Phi x\|_2^2 - \|x\|_2^2| \geq \varepsilon \|x\|_2^2) \leq 2e^{-m\varepsilon^2/K^*}$$

con $K^* = \frac{2}{1-\log(2)} \approx 6,52$.

Para llegar a ese teorema primero probaremos la Concentración de Medida para vectores aleatorios. Además, como podemos separar la desigualdad en una cota superior y una inferior, las probaremos por separado.

Teorema (Ver 4.12). *Sea $X = [X_1, \dots, X_m]$, donde los $\{X_i\}_{i=1}^m$ son i.i.d. con $X_i \sim \text{SSub}(\sigma^2)$, entonces $\mathbb{E}[\|X\|_2^2] = m\sigma^2$. Además, si fijamos $\beta_{\text{máx}} > 1$, existirá una constante $K^* \geq 6$ dependiendo solo de $\beta_{\text{máx}}$ tal que se cumple*

$$\mathbb{P}(\|X\|_2^2 \leq \alpha m\sigma^2) \leq e^{-m(1-\alpha)^2/K^*}, \text{ para cualquier } \alpha \in (0, 1),$$

y

$$\mathbb{P}(\|X\|_2^2 \geq \beta m\sigma^2) \leq e^{-m(\beta-1)^2/K^*}, \text{ para cualquier } \beta \in [1, \beta_{\text{máx}}].$$

Durante la prueba del teorema anterior observaremos que la cota superior, es decir, la desigualdad del β se puede probar directamente para Subgaussianas. De hecho, utilizando la caracterización de las variables Subgaussianas (ver Teorema 4.16) que podemos encontrar en [15] podremos probar el recíproco, que si una variable aleatoria cumple la cota superior de la Concentración de Medida será una Subgaussiana (ver Teorema 4.17).

También seremos capaces de probar el conocido Lema de Johnson-Lindenstrauss como un simple corolario de la Concentración de Medida para Estrictamente Subgaussianas. En concreto, utilizando argumentos similares a la prueba de RIP, demostraremos la siguiente versión.

Teorema (Ver 4.18). *Dado $0 < \varepsilon < 1$. Para todo conjunto de puntos Q en \mathbb{R}^N , si m es un entero positivo tal que $m > 3K^* \log(|Q|)/\varepsilon^2$, donde $K^* = \frac{2}{1-\log(2)} \approx 6,52$, entonces existe una aplicación lineal $f : \mathbb{R}^N \rightarrow \mathbb{R}^m$ tal que*

$$(1 - \varepsilon)\|u - v\|_2^2 \leq \|f(u) - f(v)\|_2^2 \leq (1 + \varepsilon)\|u - v\|_2^2, \forall u, v \in Q.$$

Acabaremos el capítulo viendo algunos ejemplos de distribuciones que son Estrictamente Subgaussianas y que, por tanto, podremos utilizar para generar las matrices aleatorias de medición. Por ejemplo, si $X \sim \text{Normal}(0, \sigma^2)$ tendremos $X \sim \text{SSub}(\sigma^2)$ o si $X \sim \text{Unif}(-a, a)$, con $a > 0$, tendremos $X \sim \text{SSub}(\frac{a^2}{3})$.

Por último, en el Capítulo 5 nos centraremos en la Minimización ℓ_1 pero de forma práctica. Utilizaremos el paquete ℓ_1 -MAGIC [8] que implementa este método en Octave [4] y lo pondremos a prueba. Con el mismo ejemplo que usábamos con el Método de los Coeficientes de Fourier comprobaremos que la Minimización ℓ_1 sí que es estable y robusta. Además, daremos una idea de los pasos a seguir para la reconstrucción de imágenes con este método.

Capítulo 2

Reconstrucción de señales sparse

Recordamos que observamos $y = \Phi x$, donde $x \in \mathbb{R}^N$ es la señal que queremos reconstruir, $y \in \mathbb{R}^m$ son las mediciones disponibles y Φ es una matriz $m \times N$ conocida. El caso que nos interesa es el indeterminado, es decir, con $m < N$, y nos preguntamos si es posible reconstruir x con buena precisión.

Durante este capítulo estudiaremos algunos métodos de reconstrucción. El primero de ellos, que es también el más elemental, nos servirá para motivar el problema, aunque como veremos no será factible en la práctica. A este método se le conoce como Minimización ℓ_0 y lo desarrollaremos en la siguiente sección, siguiendo el Capítulo 2 del libro [11].

1. Minimización ℓ_0

Denotaremos por $\|x\|_0$ el número de entradas distintas de cero del vector x , es decir,

$$\|x\|_0 := \text{card}(\text{Sop } x).$$

Recordamos que un vector x se dice *s-sparse*, denotado $x \in \Sigma_s$, precisamente cuando $\|x\|_0 \leq s$. Dado nuestro vector de mediciones, $y = \Phi x \in \mathbb{R}^m$, consideramos el siguiente problema de optimización

$$(P_0) : \left[\arg \min_{v \in \mathbb{R}^N} \|v\|_0 \text{ sujeto a } \Phi v = y \right],$$

donde $\arg \min_{v \in S} \|v\|_0 := \{v \in S : \|v\|_0 \leq \|u\|_0, \forall u \in S\}$, para $S \subseteq \mathbb{R}^N$.

Observamos que el conjunto $\{v \in \mathbb{R}^N : \Phi v = y\} = x + \ker \Phi$, por tanto, este problema busca los vectores de soporte mínimo, los más *sparse* posible, dentro de ese espacio afín. Nuestro objetivo será encontrar una matriz Φ adecuada de modo que la solución única del problema sea el vector x .

Lo primero que vamos a probar es que el número mínimo de mediciones necesarias para reconstruir de manera exacta cualquier señal *s-sparse* mediante la Minimización ℓ_0 es $2s$. Intuitivamente esto sugiere que para recuperar una señal *s-sparse* nos hace falta conocer tanto el valor de sus s componentes no nulas como las s posiciones de estas componentes.

Antes de comenzar nos hará falta introducir un poco de notación. Para un subconjunto de índices $T \subset [N] = \{1, \dots, N\}$ y una matriz Φ de tamaño $m \times N$ denotaremos por Φ_T la submatriz, de tamaño $m \times \text{card}(T)$, compuesta sólo por las columnas con índice en T .

Para un vector $x \in \mathbb{R}^N$, x_T denotará según el contexto, o bien el vector perteneciente a $\mathbb{R}^{\text{card}(T)}$ compuesto únicamente por las componentes con índice en T , o bien el vector perteneciente a \mathbb{R}^N igual a x en los índices de T e igual a cero en el resto.

Por ejemplo, si $x = (4, 5, 6) \in \mathbb{R}^3$ y $T = \{1, 3\}$, entonces según nos interese x_T será

$$x_T = (4, 6) \in \mathbb{R}^2 \quad \text{o} \quad x_T = (4, 0, 6) \in \mathbb{R}^3.$$

Para este teorema usaremos la primera interpretación.

Teorema 2.1. *Sea una matriz $\Phi \in \mathbb{R}^{m \times N}$, las siguientes propiedades son equivalentes:*

- (a) *Para cada vector $x \in \Sigma_s$, el problema (P_0) tiene como solución única el vector x .*
- (b) *El espacio nulo de la matriz, $\ker \Phi$, no contiene ningún vector $2s$ -sparse aparte del vector cero, es decir, $\ker \Phi \cap \Sigma_{2s} = \{0\}$.*
- (c) *Para cualquier conjunto de índices $T \subset [N]$ con $\text{card}(T) \leq 2s$, la función lineal asociada a la submatriz Φ_T , que va de $\mathbb{R}^{\text{card}(T)}$ a \mathbb{R}^m es inyectiva.*
- (d) *Cualquier subconjunto de $2s$ columnas de Φ es linealmente independiente.*

Demostración. Empezaremos viendo que (a) y (b) son equivalentes.

- $(a) \implies (b)$ Sea $v \in \ker \Phi \cap \Sigma_{2s}$. Tenemos que ver que $v = 0$.

Como $v \in \Sigma_{2s}$, podemos escribir $v = x - z$ con $x, z \in \Sigma_s$ y con $\text{Sop } x \cap \text{Sop } z = \emptyset$. Así, como $v \in \ker \Phi$, tendremos $z = x - v \in x + \ker \Phi$ y por la unicidad de soluciones de (P_0) concluimos que $z = x$ y, por tanto, $v = 0$.

- $(b) \implies (a)$ Sean ahora $x, z \in \Sigma_s$ tales que $z \in x + \ker \Phi$. Tenemos que ver que $x = z$.

Como $x, z \in \Sigma_s$, $x - z$ será $2s$ -sparse y además $x - z \in \ker \Phi$, luego $x - z \in \ker \Phi \cap \Sigma_{2s}$. Por tanto, como por hipótesis $\ker \Phi \cap \Sigma_{2s} = \{0\}$, $x - z = 0$ y entonces $x = z$.

Ahora, probaremos la equivalencia de (b), (c) y (d) de manera cíclica.

- $(b) \implies (c)$ Sea $T \subset [N]$ con $\text{card}(T) \leq 2s$. Vamos a ver que $\ker \Phi_T = \{0\}$.

Sea $v \in \ker \Phi_T$. Como $v \in \mathbb{R}^{\text{card}(T)}$, podemos extender el vector con ceros fuera de T . Es decir, tomamos $x \in \mathbb{R}^N$ con $\text{Sop } x = T$ tal que $x_T = v$. Así, como $\text{card}(T) \leq 2s$, $x \in \Sigma_{2s}$ y también $0 = \Phi_T v = \Phi_T x_T = \Phi x$. Por hipótesis, $\ker \Phi \cap \Sigma_{2s} = \{0\}$ luego $x = 0$ y $v = x_T = 0$. Por tanto, $\ker \Phi_T = \{0\}$ y la función lineal asociada a la matriz será inyectiva.

- $(c) \implies (d)$ Trivial, por álgebra lineal.

- $(d) \implies (b)$ Sea $v \in \ker \Phi \cap \Sigma_{2s}$. Tenemos que ver que $v = 0$.

Como $v \in \Sigma_{2s}$, podemos tomar $T \subset [N]$ con $\text{card}(T) = 2s$ tal que $\text{Sop } v \subseteq T$. Por hipótesis, las columnas de la submatriz Φ_T son linealmente independientes, luego $\ker \Phi_T = \{0\}$. Ahora, como $v \in \ker \Phi$, $v_T \in \ker \Phi_T$ y, por tanto, $v_T = 0$, lo que implica que $v = 0$. □

Si fuéramos capaces de recuperar de manera exacta cualquier señal s -sparse a partir de las mediciones $y = \Phi x \in \mathbb{R}^m$ se cumpliría (a) y, al ser equivalente, también (d). De esta manera, como cualquier subconjunto de $2s$ columnas de Φ será linealmente independiente, tendremos $\text{rank}(\Phi) \geq 2s$. Por otro lado, la matriz Φ es de tamaño $m \times N$ con $m < N$ por lo que $\text{rank}(\Phi) \leq m$. Juntando ambas relaciones llegamos a que si queremos reconstruir cualquier señal s -sparse nos harán falta al menos $2s$ mediciones.

De hecho, el siguiente teorema nos permite probar que $2s$ mediciones son suficientes.

Teorema 2.2. *Para cualquier entero $N \geq 2s$, existe una matriz Φ de tamaño $2s \times N$ tal que cualquier vector $x \in \Sigma_s$ se puede recuperar a partir de su vector de mediciones $y = \Phi x \in \mathbb{R}^{2s}$ como solución del problema (P_0) .*

Demostración. Tomamos N valores cualesquiera tales que $t_N > \dots > t_2 > t_1 > 0$ y consideramos la matriz Φ de tamaño $2s \times N$ definida como

$$\Phi = \begin{bmatrix} 1 & 1 & \dots & 1 \\ t_1 & t_2 & \dots & t_N \\ \vdots & \vdots & \ddots & \vdots \\ t_1^{2s-1} & t_2^{2s-1} & \dots & t_N^{2s-1} \end{bmatrix}.$$

Para cualquier conjunto de índices $T = \{j_1 < \dots < j_{2s}\}$ de cardinal $2s$, la submatriz cuadrada Φ_T será la transpuesta de una matriz de Vandermonde, y por tanto, como se puede ver en el Apéndice A, cumplirá

$$\det(\Phi_T) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ t_{j_1} & t_{j_2} & \dots & t_{j_{2s}} \\ \vdots & \vdots & \ddots & \vdots \\ t_{j_1}^{2s-1} & t_{j_2}^{2s-1} & \dots & t_{j_{2s}}^{2s-1} \end{vmatrix} = \prod_{1 \leq k < l \leq 2s} (t_{j_l} - t_{j_k}) > 0.$$

Luego, Φ_T será invertible y sabremos que sus columnas serán linealmente independientes. Se cumplirá, así, la propiedad (d) del teorema anterior y, entonces, cualquier vector $x \in \Sigma_s$ se podrá recuperar como la única solución de (P_0) . □

Realmente no es necesario que los números t_1, \dots, t_N sean positivos ni reales mientras se cumpla que $\det(\Phi_T) \neq 0$. Por ejemplo, podríamos tomar $t_l = e^{2\pi i(l-1)/N} \in \mathbb{C}$ para $l \in [N]$ y tendríamos la matriz

$$\Phi = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & e^{2\pi i/N} & e^{2\pi i 2/N} & \dots & e^{2\pi i(N-1)/N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{2\pi i(2s-1)/N} & e^{2\pi i(2s-1)2/N} & \dots & e^{2\pi i(2s-1)(N-1)/N} \end{bmatrix},$$

que seguiría cumpliendo (d) y, si admitimos la validez del Teorema 2.1 para vectores en \mathbb{C}^N , nos serviría para la reconstrucción. De hecho, como veremos más adelante, esta matriz se puede utilizar para representar la transformada de Fourier discreta y la usaremos para diseñar un método de reconstrucción alternativo a partir de los coeficientes de la transformada de la señal.

Aún así, aunque acabamos de ver que teóricamente se podría reconstruir cualquier señal s -sparse con tan sólo $2s$ mediciones lo cierto es que en la práctica tratar de resolver el problema (P_0) no es viable.

Podríamos pensar en aprovechar que sabemos que la solución será s -sparse y tratar de encontrarla resolviendo sistemas de ecuaciones $\Phi_T z = y$ donde T es un conjunto de índices de cardinal s . Esta vez la matriz tendrá tamaño $2s \times s$ y en caso de encontrar una solución $z \in \mathbb{R}^s$ podremos extenderla con ceros fuera de los índices de T para reconstruir nuestra señal $x \in \mathbb{R}^N$. Pero el problema está en que la cantidad de conjuntos de índices, $\binom{N}{s}$, que hay que probar es enorme.

Por ejemplo, si $N = 1000$ y $s = 10$ el número de sistemas de ecuaciones que tendríamos que resolver sería

$$\binom{1000}{10} \geq \left(\frac{1000}{10}\right)^{10} = 10^{20}.$$

Incluso si fuéramos capaces de resolver cada uno en 10^{-10} segundos, tardaríamos 10^{10} segundos (más de 300 años) en resolver el problema (P_0) .

De hecho, en la Sección 2.3 del libro [11], además de encontrarse el ejemplo que acabamos de ver, se prueba que el problema (P_0) es NP-Hard en general. Con esto queremos decir que si las clases de complejidad de problemas NP y P son distintas (como afirma la famosa conjetura $NP \neq P$) entonces no existirá un algoritmo capaz de resolver

$$\arg \min_{v \in \mathbb{R}^N} \|v\|_0 \text{ sujeto a } \Phi v = y,$$

para una matriz $\Phi \in \mathbb{R}^{m \times N}$ y un vector $y \in \mathbb{R}^m$ cualesquiera en tiempo polinomial.

No obstante, nuestro problema concreto es algo más restrictivo pues estamos trabajando con una matriz Φ conocida de antemano y nuestro vector de mediciones sabemos que es $y = \Phi x$, con x s -sparse. Así, en la siguiente sección veremos que en estas condiciones sí es posible proporcionar un método de reconstrucción en tiempo realista, con $m = 2s$ mediciones, basado en la matriz de Fourier. Aunque, como veremos, este método no va a ser ni estable ni robusto.

2. Método de los Coeficientes de Fourier

En esta sección trabajaremos con señales discretas y finitas pero no necesariamente reales, $x \in \mathbb{C}^N$. Cuando x es s -sparse lo podemos interpretar como una función x definida de $\{0, 1, \dots, N-1\}$ a \mathbb{C} con soporte igual a un conjunto $T \subset \{0, 1, \dots, N-1\}$ de cardinal s .

Para la siguiente demostración nos hará falta conocer la transformada de Fourier y la convolución discretas. Por ello, en el Apéndice B podemos encontrar un pequeño resumen con sus definiciones y un par de propiedades que usaremos.

Teorema 2.3. *Para cada $N \geq 2s$, existe un procedimiento práctico para la reconstrucción de cualquier vector s -sparse a partir de los primeros $2s$ coeficientes de su transformada de Fourier discreta.*

Demostración. Sea $x \in \mathbb{C}^N$ nuestra señal s -sparse. Suponemos que hemos podido medir los primeros $2s$ coeficientes de su transformada de Fourier discreta, $\hat{x}(0), \dots, \hat{x}(2s-1)$,

$$\text{donde } \hat{x}(j) = \sum_{k=0}^{N-1} x(k) e^{-2\pi i j k / N}, \quad 0 \leq j < N.$$

Sabemos que el soporte de x será un conjunto $T \subset \{0, 1, \dots, N-1\}$ de cardinal s . Podemos definir el siguiente polinomio trigonométrico de orden s ,

$$p(t) := \frac{1}{N} \prod_{k \in T} (1 - e^{-2\pi i k / N} e^{2\pi i t / N}),$$

cuyas raíces son exactamente los índices $t \in T$. No obstante, como en principio el conjunto T es desconocido, el objetivo de la prueba será hallar el polinomio p , o mejor dicho, su transformada \hat{p} , para estudiar sus raíces y determinar, así, quién es T .

Como la señal x tiene soporte T valdrá cero en su complementario y, por tanto, el producto $p(t)x(t) = 0$, para todo $0 \leq t < N$. Así, $\widehat{p \cdot x} = 0$ y usando la propiedad de la convolución de transformadas (ver Teorema B.6) tendremos $\widehat{p * \hat{x}} = N \widehat{p \cdot x} = 0$, es decir,

$$(\widehat{p * \hat{x}})(j) := \sum_{k=0}^{N-1} \hat{p}(k) \cdot \hat{x}(j - k \text{ mód } N) = 0, \quad \text{para todo } 0 \leq j < N. \quad (2.1)$$

Si consideramos p como una señal discreta, es decir, identificamos el polinomio $p(t)$ con $(p(0), p(1), \dots, p(N-1))$, entonces por la fórmula de inversión de la transformada de Fourier, sabemos que

$$p(t) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{p}(k) e^{2\pi i k t / N}$$

y esta fórmula es, de hecho, válida para todo $t \in \mathbb{R}$.

Luego $\frac{1}{N} \hat{p}(k)$ será el coeficiente del monomio $e^{2\pi i k t / N}$ de $p(t)$. Así, a partir del término independiente obtenemos que $\hat{p}(0) = 1$ y como el polinomio trigonométrico tenía orden s , entonces $\hat{p}(k) = 0$ para todo $k > s$. Sólo nos falta determinar los coeficientes $\hat{p}(1), \dots, \hat{p}(s)$.

Para ello, de las ecuaciones en (2.1), escribiremos las s que están en el rango $s \leq j \leq 2s-1$ en la forma

$$\begin{array}{ccccccc} \hat{x}(s) & + & \hat{p}(1)\hat{x}(s-1) & + & \dots & + & \hat{p}(s)\hat{x}(0) & = & 0, \\ \hat{x}(s+1) & + & \hat{p}(1)\hat{x}(s) & + & \dots & + & \hat{p}(s)\hat{x}(1) & = & 0, \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ \hat{x}(2s-1) & + & \hat{p}(1)\hat{x}(2s-2) & + & \dots & + & \hat{p}(s)\hat{x}(s-1) & = & 0. \end{array}$$

Que se traduce matricialmente al sistema

$$\begin{bmatrix} \hat{x}(s-1) & \hat{x}(s-2) & \dots & \hat{x}(0) \\ \hat{x}(s) & \hat{x}(s-1) & \dots & \hat{x}(1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{x}(2s-2) & \hat{x}(2s-3) & \dots & \hat{x}(s-1) \end{bmatrix} \begin{bmatrix} \hat{p}(1) \\ \hat{p}(2) \\ \vdots \\ \hat{p}(s) \end{bmatrix} = - \begin{bmatrix} \hat{x}(s) \\ \hat{x}(s+1) \\ \vdots \\ \hat{x}(2s-1) \end{bmatrix}.$$

Como $\hat{x}(0), \dots, \hat{x}(2s-1)$, son conocidos podemos resolver el sistema para obtener los $\hat{p}(j)$ que nos faltan. Observamos que el sistema tendrá al menos una solución pues por construcción el polinomio p cumple las ecuaciones. No obstante, no podemos asegurar que la matriz del sistema sea invertible por lo que podrían haber infinitas soluciones.

Por ejemplo, si nuestra señal fuera $x = (1, 0, \dots, 0)$ entonces su transformada sería el vector $\hat{x} = (1, 1, \dots, 1)$ y la matriz resultante no sería invertible. Aún así, podremos obtener una solución $\hat{q}(1), \dots, \hat{q}(s)$ que aunque no podemos asegurar que sea $\hat{p}(1), \dots, \hat{p}(s)$ nos servirá para calcular el conjunto T .

A los valores $\hat{q}(1), \dots, \hat{q}(s)$ conocidos les añadimos $\hat{q}(0) = 1$ y $\hat{q}(k) = 0$ para $k > s$. Para estos \hat{q} , volveremos a tener el sistema

$$(\hat{q} * \hat{x})(j) = 0, \quad \text{para todo } s \leq j \leq 2s-1.$$

Por tanto, como $N\widehat{q \cdot x} = \widehat{q} * \hat{x}$, el vector $q \cdot x$ tendrá transformada de Fourier igual a 0 en s índices consecutivos. Si escribimos esas ecuaciones en forma matricial tenemos

$$\begin{bmatrix} 1 & e^{2\pi i s/N} & e^{2\pi i 2s/N} & \dots & e^{2\pi i (N-1)s/N} \\ 1 & e^{2\pi i (s+1)/N} & e^{2\pi i 2(s+1)/N} & \dots & e^{2\pi i (N-1)(s+1)/N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{2\pi i (2s-1)/N} & e^{2\pi i 2(2s-1)/N} & \dots & e^{2\pi i (N-1)(2s-1)/N} \end{bmatrix} \begin{bmatrix} (q \cdot x)(0) \\ (q \cdot x)(1) \\ \vdots \\ (q \cdot x)(N-1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Aunque la matriz del sistema, Φ , es de tamaño $s \times N$, por su forma, cualquier submatriz de s columnas será una matriz de tipo Vandermonde construida a partir de elementos diferentes. Entonces, como veíamos en el Apéndice A, la submatriz tendrá determinante no nulo y será invertible.

Así, como x era s -sparse, $q \cdot x$ también lo será y podremos restringirnos a la submatriz de las s columnas del soporte T . Como hemos dicho esta submatriz será invertible y nos dará que $q \cdot x = 0$ en T . Por tanto, como fuera de T ya valía 0 por x , la única solución posible será la nula,

$$(q \cdot x)(j) = 0, \quad \text{para todo } 0 \leq j < N.$$

Ahora bien, sabemos que x tiene como soporte el conjunto de índices T por lo que debe ser q el que valga 0 en T . Como los valores de su transformada ya los conocemos podemos volver a usar la fórmula de inversión y tendremos

$$q(t) = \frac{1}{N} \sum_{0 \leq k < N} \hat{q}(k) e^{2\pi i kt/N} = \frac{1}{N} \sum_{k=0}^s \hat{q}(k) e^{2\pi i kt/N}.$$

Por construcción, q es un polinomio trigonométrico de grado s y, como el cardinal de T es s , las únicas raíces del polinomio deberán ser los elementos de T . De esta manera, podemos determinar el conjunto de índices T calculando las raíces de q .

De hecho, en la práctica, como las raíces serán algunos de los índices podemos obtenerlas buscando los s valores más pequeños de $|q(j)|$, $0 \leq j < N$ y en ningún momento nos hará falta construir el polinomio como tal.

Una vez ya conocemos el soporte de nuestra señal x , los valores $x(t)$, $t \in T$, los podemos obtener resolviendo el sistema de ecuaciones que nos proporciona la transformada.

Como ya conocíamos $\hat{x}(0), \dots, \hat{x}(2s-1)$, tendremos las ecuaciones

$$\hat{x}(j) = \sum_{t \in T} x(t) e^{-2\pi i j t / N}, \quad \text{para } 0 \leq j \leq 2s-1.$$

Y, denotando $T = \{t_1, \dots, t_s\}$, se podrá escribir en forma matricial como

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-2\pi i t_1 / N} & e^{-2\pi i t_2 / N} & \dots & e^{-2\pi i t_s / N} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-2\pi i (2s-1)t_1 / N} & e^{-2\pi i (2s-1)t_2 / N} & \dots & e^{-2\pi i (2s-1)t_s / N} \end{bmatrix} \begin{bmatrix} x(t_1) \\ x(t_2) \\ \vdots \\ x(t_s) \end{bmatrix} = \begin{bmatrix} \hat{x}(0) \\ \hat{x}(1) \\ \vdots \\ \hat{x}(2s-1) \end{bmatrix}.$$

Ahora el sistema es sobredeterminado ($2s \times s$) pero sabemos que la señal x cumplirá las ecuaciones por lo que al menos tendrá una solución. Además, como la matriz vuelve a tener la forma de una matriz de Vandermonde, tendrá rango máximo y la solución será única.

□

Aunque es verdad que este procedimiento resuelve el problema del elevado tiempo de computación, sigue presentando inconvenientes que lo hacen inviable en la práctica. Como vamos a ver a continuación con un par de ejemplos, el proceso ni es ni robusto respecto a errores de medición ni estable cuando la señal deja de ser exactamente *sparse*.

Para ilustrarlo trabajaremos con una señal x de tamaño $N = 500$ compuesta por $s = 20$ picos ± 1 generados aleatoriamente y el resto ceros. Usaremos la implementación del procedimiento en el programa Octave [4] que podemos encontrar en el Apéndice D y seguiremos las indicaciones del Ejercicio 2.8 del libro [11].

Vemos que efectivamente el método es capaz de reconstruir de manera exacta la señal con tan solo los primeros $2s$ coeficientes de la transformada de Fourier.

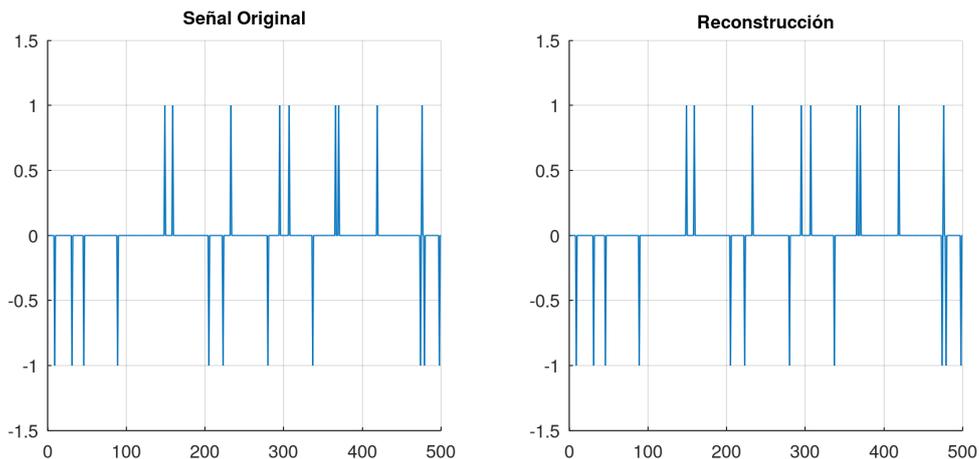


Figura 2.1: Reconstrucción exacta partir de $2s$ coeficientes de Fourier.

En cuanto a la robustez, si a la hora de obtener los coeficientes de la transformada cometiéramos algún error de medición la reconstrucción no sería exacta. Por ejemplo, si al coeficiente s -ésimo de la transformada de la señal original le sumamos simplemente 0.001, es decir, si $y(s) = \hat{x}(s) + 0.001$, obtendremos:

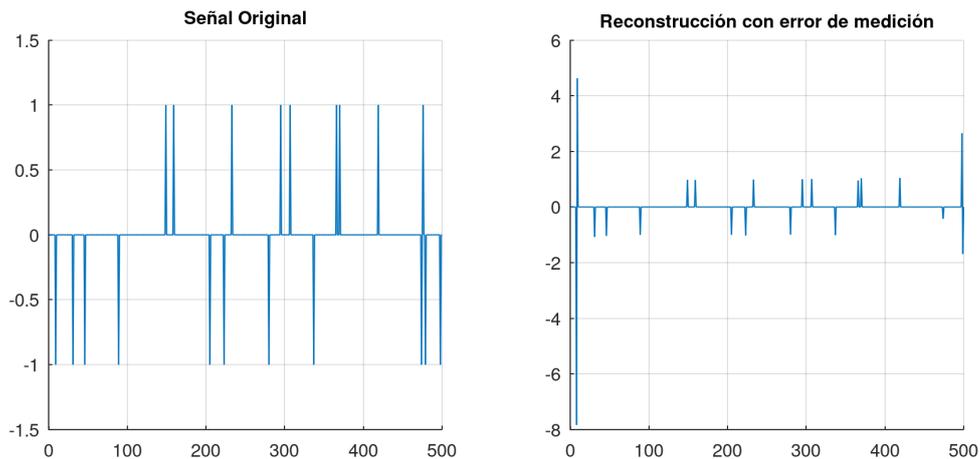


Figura 2.2: Problemas de robustez en la reconstrucción de Fourier.

Por la construcción de la señal x su norma valdrá $\|x\|_2 = \sqrt{20} \approx 4.472$ y si x^* es la señal reconstruida podemos incluso medir el error relativo $\frac{\|x - x^*\|_2}{\|x\|_2}$.

```
>> x0 = FourierMethod(N,s,y);
>> norm(x-x0)/norm(x)
ans = 4.0917
```

Este error es muy elevado teniendo en cuenta que nuestro error de medición es sólo 0.001 y que la mayoría de coeficientes de Fourier de x tienen módulos por encima de 2. En particular, el que modificamos es $\hat{x}(20) = 4$.

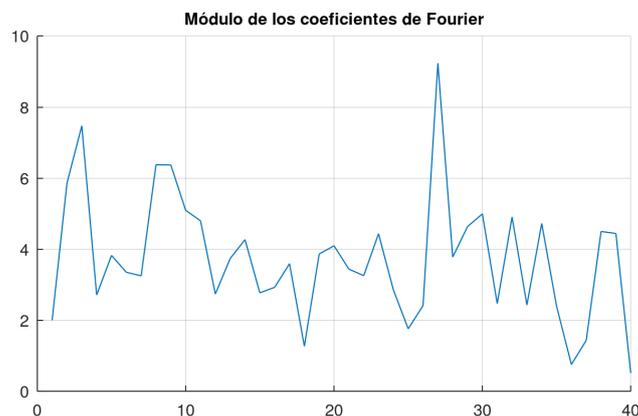


Figura 2.3: Módulo de los primeros $2s$ coeficientes de la TFD de la señal.

Para comprobar la estabilidad añadiremos a la señal original un nuevo pico, esta vez de tamaño 0.1 y veremos que ya no es capaz de recuperar la señal de manera exacta. Además, la reconstrucción varía mucho dependiendo de la posición donde añadamos el pico.

Si el pico lo añadimos en la posición 180, es decir, $x(180) = 0.1$ en vez de 0, obtendremos una reconstrucción bastante buena pero que no llega a ser exacta.

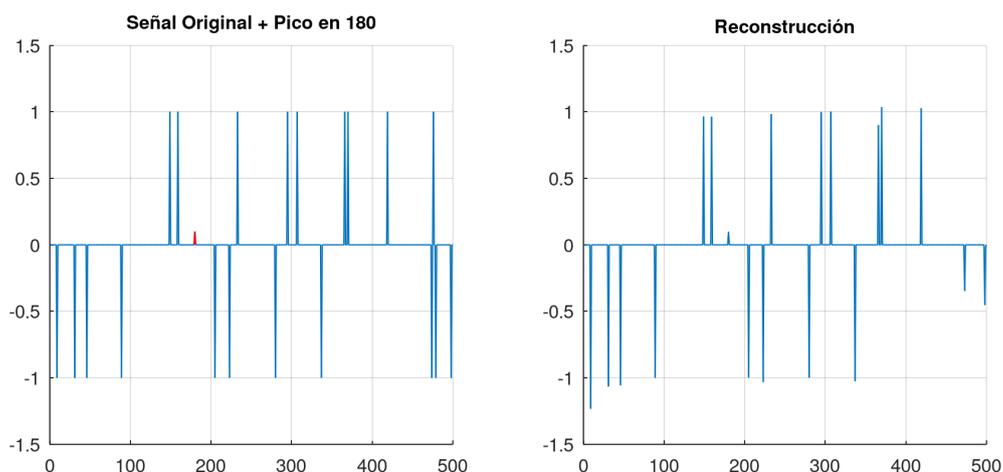


Figura 2.4: Problemas de estabilidad en la reconstrucción de Fourier (Pico en 180).

Si, en cambio, el pico lo añadimos en la posición 200 la reconstrucción será mucho peor.

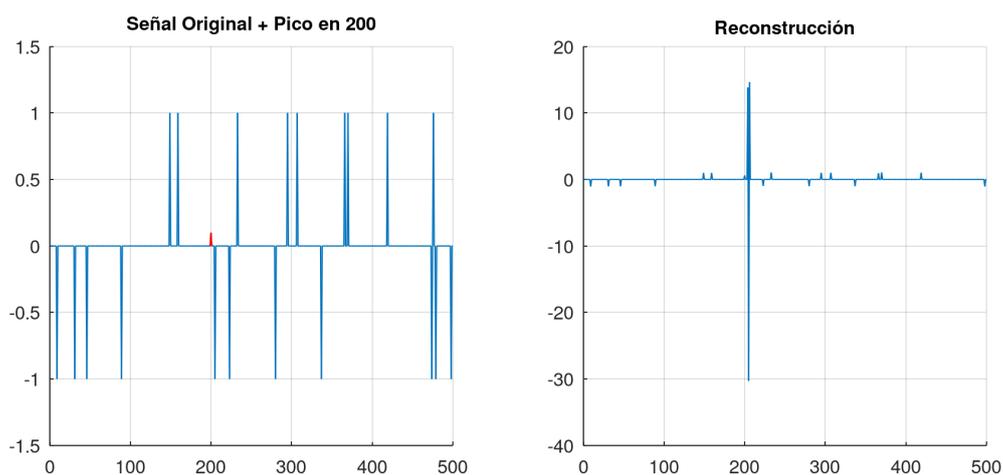


Figura 2.5: Problemas de estabilidad en la reconstrucción de Fourier (Pico en 200).

Evidentemente, esto se ve reflejado en los errores relativos. Mientras que con el pico en 180 el error relativo es 0.8632, con el pico en 200 se alcanza un error de 9.7763.

En definitiva, aunque estos métodos son llamativos por el pequeño número de mediciones que necesitan, hemos visto que no son viables en la práctica. Por eso, para acabar el capítulo, trataremos de encontrar un método de reconstrucción alternativo que aunque requiera un mayor número de mediciones sea estable, robusto y posea algoritmos de resolución eficientes.

3. Normas ℓ_p

Al estar trabajando con vectores de dimensión finita, podemos utilizar las normas ℓ_p , para $1 \leq p \leq \infty$. Así, para el espacio vectorial \mathbb{R}^d se definen las normas ℓ_p^d como:

$$\|z\|_{\ell_p^d} := \left(\sum_{i=1}^d |z_i|^p \right)^{1/p}, \quad \text{para } 1 \leq p < \infty.$$

$$\|z\|_{\ell_\infty^d} := \max_{1 \leq i \leq d} |z_i|, \quad \text{para } p = \infty.$$

Abusaremos de notación omitiendo la dimensión del espacio y denotaremos $\|z\|_p = \|z\|_{\ell_p^d}$.

Con estas funciones podemos tratar de reconstruir la señal x a partir de las soluciones de los problemas de optimización,

$$\arg \min_{v \in \mathbb{R}^N} \|v\|_p \text{ sujeto a } \Phi v = y.$$

De estas normas las más habituales son la ℓ_1 y la ℓ_2 . No obstante, como veremos en un ejemplo, la minimización con norma ℓ_2 casi nunca devuelve un vector *sparse* y, por tanto, no nos servirá para la reconstrucción de señales.

Supongamos que nuestra señal es $x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, o sea, una señal de tamaño $N = 2$ pero que es 1-*sparse*. Además, supongamos que $m = 1$ y la matriz de medición es $\Phi = \begin{pmatrix} 1 & 2 \end{pmatrix}$. Entonces nuestra única medición será $y = \Phi x = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2$. Las soluciones que vamos a buscar están entre los vectores $v \in \mathbb{R}^2$ que cumplen $y = \Phi v$. Si hacemos el cálculo, tenemos

$$2 = y = \Phi v = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = v_1 + 2v_2 \iff v_1 = 2(1 - v_2).$$

Si representamos gráficamente las soluciones observamos como la forma de las bolas en cada norma determina lo eficaz que es el método en la reconstrucción.

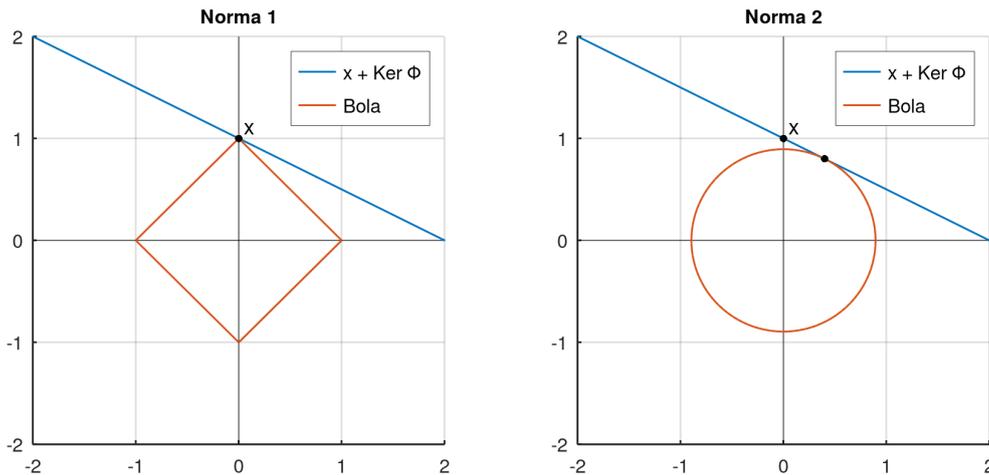


Figura 2.6: Ejemplo de los algoritmos de minimización con las normas ℓ_1 y ℓ_2 .

De hecho, si denotamos la matriz de medición $\Phi = \begin{pmatrix} a_1 & a_2 \end{pmatrix}$ se puede ver que con la norma ℓ_2 solamente se reconstruirá la señal x de manera exacta cuando $a_1 = 0$. Mientras que, con la norma ℓ_1 valdrán a_1 y a_2 cualesquiera siempre que $|a_1| < |a_2|$.

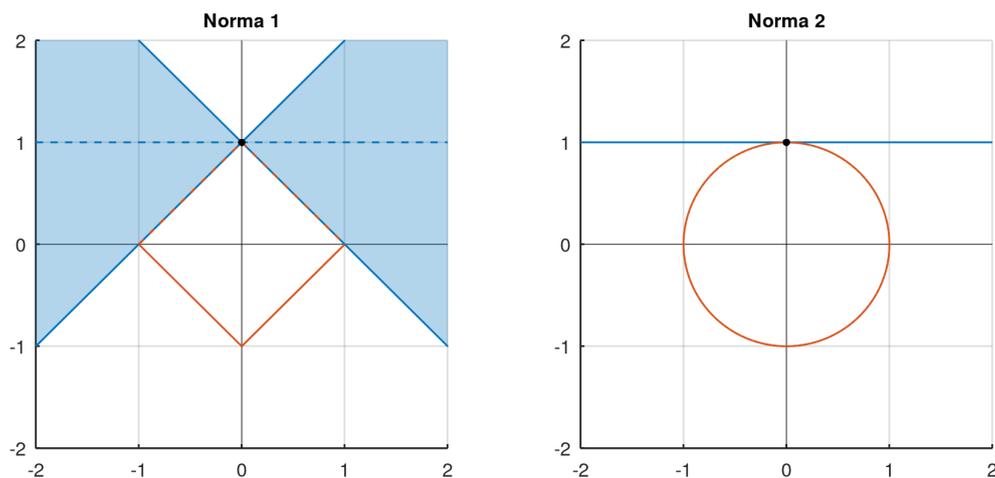


Figura 2.7: Regiones donde los algoritmos reconstruyen la señal de manera exacta.

Evidentemente, en dimensiones superiores seguirá ocurriendo lo mismo. El método de Minimización en norma ℓ_1 con alta probabilidad recuperará las señales de manera exacta pero el de la norma ℓ_2 no. Por ello, durante el desarrollo teórico que haremos en el siguiente capítulo el método de reconstrucción que utilizaremos será el de la norma ℓ_1 .

Además, la ventaja de trabajar con una norma es que será convexa y, por tanto, podremos utilizar técnicas de optimización convexa para resolver los problemas de minimización. De hecho, todo el Capítulo 15 del libro [11] está dedicado al estudio de algoritmos eficientes para resolver la Minimización en norma ℓ_1 .

Aunque nosotros en este trabajo no estudiaremos la programación explícita de algoritmos de minimización en ℓ_1 , en el Capítulo 5 sí que utilizaremos una implementación del método en Octave [4] para ponerlo a prueba con algunos ejemplos.

Capítulo 3

Compressed Sensing mediante la Minimización ℓ_1

Ahora que hemos seleccionado un método de reconstrucción vamos a analizarlo para ver qué características posee. Veremos que, escogiendo debidamente la matriz de medición, podremos asegurar con probabilidad alta que la reconstrucción será exacta, estable y robusta. Además, para este método también encontraremos el número mínimo de mediciones necesarias, que en este caso será $m = \mathcal{O}(s \log(N/s))$; ver Teorema 3.8 más adelante.

Como ya sabemos, la reconstrucción con la minimización en norma ℓ_1 se basa en resolver el siguiente problema de optimización. Dado un vector $x \in \mathbb{R}^N$, si $y = \Phi x$ buscamos las soluciones de

$$(P) : \left[\arg \min_{v \in \mathbb{R}^N} \|v\|_1 \text{ sujeto a } \Phi v = y \right].$$

Nuestro primer objetivo será probar que cualquier solución x^* del problema (P) recuperará de manera exacta la señal x si esta es suficientemente *sparse* y si la matriz Φ cumple la Propiedad de Isometría Restringida. Para ello, nos basaremos en el artículo [7].

1. La Propiedad de Isometría Restringida

Definición 3.1. Para cada entero positivo s se define la constante de isometría de una matriz Φ de tamaño $m \times N$, como el menor número $\delta_s \geq 0$ tal que

$$(1 - \delta_s)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta_s)\|x\|_2^2, \quad \forall x \in \Sigma_s. \quad (3.1)$$

Se dice que la matriz satisface la Propiedad de Isometría Restringida (RIP) de orden s si

$$0 \leq \delta_s < 1.$$

Realmente el resultado que probaremos no se centrará solamente en recuperar vectores *sparse* sino en recuperar vectores cualesquiera x de \mathbb{R}^N . Por tanto, para medir la calidad (o estabilidad) de la reconstrucción compararemos x^* con la mejor aproximación *s-sparse* de x , es decir, un vector $x_s \in \Sigma_s$ que cumple

$$\|x - x_s\|_1 = \inf_{z \in \Sigma_s} \|x - z\|_1. \quad (*)$$

Dichos vectores x_s se construyen tomando las s componentes más grandes de x (en valor absoluto) y haciendo cero el resto. Obsérvese que x_s podría no ser único (si x tiene varias componentes con el mismo valor absoluto), en cuyo caso el resultado que sigue vale para cualquier elección que cumpla (*).

Teorema 3.1 (Reconstrucción estable (sin ruido)).

Si suponemos que $\delta_{2s} < \sqrt{2} - 1$, entonces cualquier solución x^* del problema (P) cumple

$$\|x^* - x\|_2 \leq \frac{c_0}{\sqrt{s}} \|x - x_s\|_1 \quad (3.2)$$

para una constante $c_0 > 0$, que depende únicamente de δ_{2s} . En particular, si x es s -sparse, la reconstrucción será exacta.

De hecho, podemos probar este resultado para el caso más general en el que tenemos mediciones con ruido, $y = \Phi x + z$. Ahora lo que propondremos es reconstruir x a partir de las soluciones del problema

$$(P^*) : \left[\arg \min_{v \in \mathbb{R}^N} \|v\|_1 \text{ sujeto a } \|y - \Phi v\|_2 \leq \varepsilon \right],$$

donde ε es una cota superior del ruido.

Teorema 3.2 (Reconstrucción estable y robusta (con ruido)).

Si suponemos que $\delta_{2s} < \sqrt{2} - 1$ y $\|y - \Phi x\|_2 \leq \varepsilon$, entonces si x^* es solución del problema (P*) se cumple

$$\|x^* - x\|_2 \leq \frac{c_0}{\sqrt{s}} \|x - x_s\|_1 + c_1 \varepsilon \quad (3.3)$$

para constantes $c_0, c_1 > 0$, que dependen de δ_{2s} .

Para probarlo necesitaremos el siguiente lema:

Lema 3.3. Para cualesquiera $x, x' \in \mathbb{R}^N$ con soportes disjuntos $T, T' \subset \{1, \dots, N\}$ tales que $\text{card}(T) \leq s$ y $\text{card}(T') \leq s'$ se cumple

$$|\langle \Phi x, \Phi x' \rangle| \leq \delta_{s+s'} \|x\|_2 \|x'\|_2. \quad (3.4)$$

Demostración. Empezaremos suponiendo que tanto x como x' son vectores unitarios.

Así, al tener soportes disjuntos, tendremos $\|x \pm x'\|_2^2 = \|x\|_2^2 + \|x'\|_2^2 = 2$.

Además, $x \pm x'$ serán vectores $(s + s')$ -sparse y por la definición de constante de isometría,

$$(1 - \delta_{s+s'}) \|x \pm x'\|_2^2 \leq \|\Phi(x \pm x')\|_2^2 \leq (1 + \delta_{s+s'}) \|x \pm x'\|_2^2,$$

con lo que,

$$2(1 - \delta_{s+s'}) \leq \|\Phi x \pm \Phi x'\|_2^2 \leq 2(1 + \delta_{s+s'}).$$

Luego, tanto $\|\Phi x + \Phi x'\|_2^2$ como $\|\Phi x - \Phi x'\|_2^2$ están en un intervalo de longitud $4\delta_{s+s'}$ y, por tanto, su diferencia cumple

$$-4\delta_{s+s'} \leq \|\Phi x + \Phi x'\|_2^2 - \|\Phi x - \Phi x'\|_2^2 \leq 4\delta_{s+s'}.$$

Entonces, por la ley del paralelogramo, tenemos

$$|\langle \Phi x, \Phi x' \rangle| = \frac{1}{4} \left| \|\Phi x + \Phi x'\|_2^2 - \|\Phi x - \Phi x'\|_2^2 \right| \leq \delta_{s+s'}.$$

Si x y x' no fueran unitarios, y suponiendo que ninguno es nulo haríamos

$$|\langle \Phi x, \Phi x' \rangle| = \|x\|_2 \|x'\|_2 \left| \left\langle \Phi \frac{x}{\|x\|_2}, \Phi \frac{x'}{\|x'\|_2} \right\rangle \right| \leq \delta_{s+s'} \|x\|_2 \|x'\|_2.$$

□

Ahora ya podemos probar el Teorema 3.2 de reconstrucción con ruido.

Dem. del Teorema 3.2. Recordamos, por un lado, que por hipótesis, $\|y - \Phi x\|_2 \leq \varepsilon$ y, por otro, que x^* es solución de (P^*) luego $\|y - \Phi x^*\|_2 \leq \varepsilon$. Así, tendremos

$$\|\Phi(x^* - x)\|_2 \leq \|\Phi x^* - y\|_2 + \|y - \Phi x\|_2 \leq 2\varepsilon. \quad (3.5)$$

Recordamos que dado un conjunto de índices $T \subset \{1, \dots, N\}$ y un vector $v \in \mathbb{R}^N$, denotábamos v_T al vector de \mathbb{R}^N igual a v en los índices de T y 0 en el resto.

Escribimos $x^* = x + h$ y descomponemos h en una suma de vectores s -sparse, h_{T_0}, h_{T_1}, \dots donde: T_0 son los índices de las s componentes más grandes de x .

T_1 son los índices de las s componentes más grandes de $h_{T_0^c}$.

T_2 son los índices de las siguientes s componentes más grandes de $h_{T_1^c}$.

Y así sucesivamente hasta no poder crear más conjuntos.

Observamos que todos los conjuntos tendrán cardinal s , excepto quizás el último.

La prueba consiste en acotar $\|h_{T_0 \cup T_1}\|_2$ y $\|h_{(T_0 \cup T_1)^c}\|_2$.

$$\text{Dado } j \geq 2, \quad \|h_{T_j}\|_2^2 = \sum_{i \in T_j} |h_i|^2 \leq \sum_{i \in T_j} \max_{k \in T_j} |h_k|^2 = \|h_{T_j}\|_\infty^2 \sum_{i \in T_j} 1 \stackrel{|T_j| \leq s}{\leq} s \|h_{T_j}\|_\infty^2.$$

Como, por construcción de los conjuntos T_j , cualquier componente de $h_{T_{j-1}}$ será mayor en módulo que todas las componentes de h_{T_j} , tendremos

$$\|h_{T_j}\|_2 \leq \sqrt{s} \|h_{T_j}\|_\infty = \frac{\sqrt{s}}{s} \sum_{i \in T_{j-1}} \|h_{T_j}\|_\infty \leq \frac{1}{\sqrt{s}} \|h_{T_{j-1}}\|_1.$$

Así, si sumamos para $j \geq 2$,

$$\sum_{j \geq 2} \|h_{T_j}\|_2 \leq \frac{1}{\sqrt{s}} (\|h_{T_1}\|_1 + \|h_{T_2}\|_1 + \dots) = \frac{1}{\sqrt{s}} \|h_{T_0^c}\|_1.$$

Se tiene entonces

$$\|h_{(T_0 \cup T_1)^c}\|_2 = \left\| \sum_{j \geq 2} h_{T_j} \right\|_2 \leq \sum_{j \geq 2} \|h_{T_j}\|_2 \leq \frac{1}{\sqrt{s}} \|h_{T_0^c}\|_1. \quad (3.6)$$

Recordamos que $x^* = x + h$ es la solución de (P^*) : $\left[\arg \min_{v \in \mathbb{R}^N} \|v\|_1 \text{ sujeto a } \|y - \Phi v\|_2 \leq \varepsilon \right]$

con lo que $\|x\|_1 \geq \|x^*\|_1 = \|x + h\|_1 = \sum_{i \in T_0} |x_i + h_i| + \sum_{i \in T_0^c} |x_i + h_i|.$

Usando la propiedad del valor absoluto, $|x_i + h_i| \geq \left| |x_i| - |h_i| \right|$, nos quedará

$$\|x\|_1 \geq \sum_{i \in T_0} |x_i| - \sum_{i \in T_0} |h_i| + \sum_{i \in T_0^c} |h_i| - \sum_{i \in T_0^c} |x_i| = \|x_{T_0}\|_1 - \|h_{T_0}\|_1 + \|h_{T_0^c}\|_1 - \|x_{T_0^c}\|_1.$$

Lo que queremos es acotar $\|h_{T_0^c}\|_1$ así que despejando ese término tenemos

$$\|h_{T_0^c}\|_1 \leq \|h_{T_0}\|_1 + \underbrace{\|x\|_1 - \|x_{T_0}\|_1}_{\|x_{T_0^c}\|_1} + \|x_{T_0^c}\|_1 = \|h_{T_0}\|_1 + 2\|x_{T_0^c}\|_1.$$

Sustituyendo en la ecuación (3.6) llegamos a

$$\|h_{(T_0 \cup T_1)^c}\|_2 \leq \frac{1}{\sqrt{s}} \|h_{T_0^c}\|_1 \leq \frac{1}{\sqrt{s}} \|h_{T_0}\|_1 + \frac{2}{\sqrt{s}} \|x_{T_0^c}\|_1. \quad (3.7)$$

Recordamos que T_0 son los índices de las s componentes más grandes de x , luego $x_{T_0} = x_s$ y entonces $x_{T_0^c} = x - x_s$. Usando la desigualdad de Cauchy-Schwarz, podemos relacionar la norma ℓ_1 de h_{T_0} con su norma ℓ_2 , $\|h_{T_0}\|_1 \leq \sqrt{s} \|h_{T_0}\|_2$.

Así, de la ecuación (3.7) pasamos a

$$\|h_{(T_0 \cup T_1)^c}\|_2 \leq \|h_{T_0}\|_2 + \frac{2}{\sqrt{s}} \|x - x_s\|_1 \leq \|h_{T_0 \cup T_1}\|_2 + \frac{2}{\sqrt{s}} \|x - x_s\|_1. \quad (3.8)$$

Ahora vamos a acotar $\|h_{T_0 \cup T_1}\|_2$. Observamos que $\Phi h_{T_0 \cup T_1} = \Phi h - \sum_{j \geq 2} \Phi h_{T_j}$, por lo que podemos calcular su norma haciendo

$$\|\Phi h_{T_0 \cup T_1}\|_2^2 = \langle \Phi h_{T_0 \cup T_1}, \Phi h - \sum_{j \geq 2} \Phi h_{T_j} \rangle. \quad (3.9)$$

Haremos primero el cálculo de $\langle \Phi h_{T_0 \cup T_1}, \Phi h \rangle$ y después el de $\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \rangle$.

Volviendo a utilizar la desigualdad de Cauchy-Schwarz,

$$|\langle \Phi h_{T_0 \cup T_1}, \Phi h \rangle| \leq \|\Phi h_{T_0 \cup T_1}\|_2 \|\Phi h\|_2.$$

Como T_0 y T_1 son conjuntos de índices disjuntos con cardinal $\leq s$, el vector $h_{T_0 \cup T_1}$ será $2s$ -sparse y por la definición de la constante de isometría obtenemos

$$(1 - \delta_{2s}) \|h_{T_0 \cup T_1}\|_2^2 \leq \|\Phi h_{T_0 \cup T_1}\|_2^2 \leq (1 + \delta_{2s}) \|h_{T_0 \cup T_1}\|_2^2.$$

Y la desigualdad de la derecha se queda en

$$\|\Phi h_{T_0 \cup T_1}\|_2 \leq \sqrt{1 + \delta_{2s}} \|h_{T_0 \cup T_1}\|_2.$$

Recordamos que, por (3.5), $\|\Phi h\|_2 = \|\Phi(x^* - x)\|_2 \leq 2\varepsilon$, con lo que llegamos a

$$|\langle \Phi h_{T_0 \cup T_1}, \Phi h \rangle| \leq \|\Phi h_{T_0 \cup T_1}\|_2 \|\Phi h\|_2 \leq 2\varepsilon \sqrt{1 + \delta_{2s}} \|h_{T_0 \cup T_1}\|_2. \quad (3.10)$$

Para el cálculo de $\langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \rangle$ hacemos uso del Lema 3.3 y obtenemos

$$|\langle \Phi h_{T_0}, \Phi h_{T_j} \rangle| \leq \delta_{2s} \|h_{T_0}\|_2 \|h_{T_j}\|_2 \quad \text{y} \quad |\langle \Phi h_{T_1}, \Phi h_{T_j} \rangle| \leq \delta_{2s} \|h_{T_1}\|_2 \|h_{T_j}\|_2.$$

Para cualesquiera a y b se cumple $a + b \leq \sqrt{2}\sqrt{a^2 + b^2}$. Aplicando este resultado con $a = \|h_{T_0}\|_2$ y $b = \|h_{T_1}\|_2$ tenemos

$$\begin{aligned} \|h_{T_0}\|_2 + \|h_{T_1}\|_2 &\leq \sqrt{2}\sqrt{\|h_{T_0}\|_2^2 + \|h_{T_1}\|_2^2} = \sqrt{2}\sqrt{\sum_{i \in T_0} |h_i|^2 + \sum_{i \in T_1} |h_i|^2} = \\ &= \sqrt{2}\sqrt{\sum_{i \in T_0 \cup T_1} |h_i|^2} = \sqrt{2}\|h_{T_0 \cup T_1}\|_2. \end{aligned}$$

Con todo esto

$$\begin{aligned} \left| \langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \rangle \right| &\leq \left| \langle \Phi h_{T_0}, \sum_{j \geq 2} \Phi h_{T_j} \rangle \right| + \left| \langle \Phi h_{T_1}, \sum_{j \geq 2} \Phi h_{T_j} \rangle \right| \leq \\ &\leq \delta_{2s} \|h_{T_0}\|_2 \sum_{j \geq 2} \|h_{T_j}\|_2 + \delta_{2s} \|h_{T_1}\|_2 \sum_{j \geq 2} \|h_{T_j}\|_2 = \\ &= \delta_{2s} (\|h_{T_0}\|_2 + \|h_{T_1}\|_2) \sum_{j \geq 2} \|h_{T_j}\|_2 \leq \delta_{2s} \sqrt{2} \|h_{T_0 \cup T_1}\|_2 \sum_{j \geq 2} \|h_{T_j}\|_2. \end{aligned} \quad (3.11)$$

Por tanto, juntando (3.9), (3.10) y (3.11) obtenemos

$$\begin{aligned} \|\Phi h_{T_0 \cup T_1}\|_2^2 &\leq \left| \langle \Phi h_{T_0 \cup T_1}, \Phi h \rangle \right| + \left| \langle \Phi h_{T_0 \cup T_1}, \sum_{j \geq 2} \Phi h_{T_j} \rangle \right| \leq \\ &\leq \|h_{T_0 \cup T_1}\|_2 \left(2\varepsilon \sqrt{1 + \delta_{2s}} + \sqrt{2} \delta_{2s} \sum_{j \geq 2} \|h_{T_j}\|_2 \right). \end{aligned}$$

Que junto con la cota inferior de la constante de isometría nos da

$$(1 - \delta_{2s}) \|h_{T_0 \cup T_1}\|_2^2 \leq \|\Phi h_{T_0 \cup T_1}\|_2^2 \leq \|h_{T_0 \cup T_1}\|_2 \left(2\varepsilon \sqrt{1 + \delta_{2s}} + \sqrt{2} \delta_{2s} \sum_{j \geq 2} \|h_{T_j}\|_2 \right).$$

El objetivo era acotar $\|h_{T_0 \cup T_1}\|_2$ así que hacemos

$$\begin{aligned} \|h_{T_0 \cup T_1}\|_2 &\leq \overbrace{\left(\frac{2\sqrt{1 + \delta_{2s}}}{1 - \delta_{2s}} \right)}^{\alpha > 0} \varepsilon + \overbrace{\left(\frac{\sqrt{2} \delta_{2s}}{1 - \delta_{2s}} \right)}^{\rho > 0} \sum_{j \geq 2} \|h_{T_j}\|_2 \stackrel{(3.6)}{\leq} \\ &\leq \alpha \varepsilon + \rho \frac{1}{\sqrt{s}} \|h_{T_0^c}\|_1 \stackrel{(3.7)}{\leq} \alpha \varepsilon + \rho \frac{1}{\sqrt{s}} \|h_{T_0}\|_1 + \rho \frac{2}{\sqrt{s}} \|x - x_s\|_1 \leq \\ &\leq \alpha \varepsilon + \rho \|h_{T_0 \cup T_1}\|_2 + \rho \frac{2}{\sqrt{s}} \|x - x_s\|_1. \end{aligned}$$

Si $\rho < 1$, juntando los $\|h_{T_0 \cup T_1}\|_2$ se llega a

$$\|h_{T_0 \cup T_1}\|_2 \leq \frac{1}{1 - \rho} \left(\alpha \varepsilon + \frac{2\rho}{\sqrt{s}} \|x - x_s\|_1 \right). \quad (3.12)$$

Gracias a la hipótesis $\delta_{2s} < \sqrt{2} - 1$ podemos asegurar que $\rho < 1$.

$$1 > \rho = \frac{\sqrt{2} \delta_{2s}}{1 - \delta_{2s}} \iff 1 - \delta_{2s} > \sqrt{2} \delta_{2s} \iff \delta_{2s} < \frac{1}{\sqrt{2} + 1} = \sqrt{2} - 1$$

Finalmente, haciendo uso de las cotas (3.8) y (3.12) obtenemos lo que queríamos

$$\begin{aligned} \|x^* - x\|_2 &= \|h\|_2 \leq \|h_{T_0 \cup T_1}\|_2 + \|h_{(T_0 \cup T_1)^c}\|_2 \leq 2\|h_{T_0 \cup T_1}\|_2 + \frac{2}{\sqrt{s}}\|x - x_s\|_1 \leq \\ &\leq \underbrace{\left(2 + \frac{4\rho}{1-\rho}\right)}_{c_0 > 0} \frac{1}{\sqrt{s}}\|x - x_s\|_1 + \underbrace{\left(\frac{2\alpha}{1-\rho}\right)}_{c_1 > 0} \varepsilon = \frac{c_0}{\sqrt{s}}\|x - x_s\|_1 + c_1\varepsilon. \end{aligned}$$

□

Una vez probado este resultado, el Teorema 3.1 de reconstrucción sin ruido es consecuencia directa tomando $\varepsilon = 0$. Para las señales $x \in \Sigma_s$ la reconstrucción es exacta pues se tiene $x = x_s$.

Por supuesto, lo que nos interesa ahora es construir matrices de medición que cumplan la propiedad RIP. En la siguiente sección, basándonos en el artículo [2], demostraremos que las matrices aleatorias generadas a partir de una distribución que cumple la llamada propiedad de Concentración de Medida cumplirán también RIP con una probabilidad alta.

2. La Propiedad de Concentración de Medida

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea r una variable aleatoria en Ω . Podemos generar matrices aleatorias, Φ , de tamaño $m \times N$ tomando sus entradas $\Phi_{i,j}$ como realizaciones independientes de la variable r . Así tendremos las matrices $\Phi(\omega)$, $\omega \in \Omega$.

Probaremos que se cumple la propiedad RIP con probabilidad alta partiendo de que las matrices aleatorias $\Phi(\omega)$ cumplen las siguientes propiedades:

1) Esperanza

Para cualquier $x \in \mathbb{R}^N$, la variable aleatoria $\|\Phi(\omega)x\|_2^2$ tiene esperanza igual a $\|x\|_2^2$,

$$\mathbb{E}(\|\Phi(\omega)x\|_2^2) = \|x\|_2^2. \quad (3.13)$$

2) Concentración de Medida

Para cualquier $x \in \mathbb{R}^N$, la variable aleatoria $\|\Phi(\omega)x\|_2^2$ está concentrada alrededor de su esperanza. Dado $0 < \varepsilon < 1$,

$$\mathbb{P}\left(\left|\|\Phi(\omega)x\|_2^2 - \|x\|_2^2\right| \geq \varepsilon\|x\|_2^2\right) \leq 2e^{-mc_0(\varepsilon)}, \quad (3.14)$$

donde la probabilidad es considerada sobre todas las matrices aleatorias $m \times N$ y $c_0(\varepsilon)$ es una constante que depende solo de ε y tal que $\forall \varepsilon \in (0, 1)$, $c_0(\varepsilon) > 0$.

Ya en el Capítulo 4 nos centraremos en las matrices aleatorias generadas a partir de una distribución Subgaussiana y probaremos que cumplen estas propiedades y que, por tanto, se pueden utilizar para la reconstrucción de señales. Además, en el Apéndice C podemos encontrar un pequeño resumen de los conceptos básicos de Teoría de la Probabilidad.

Nos harán falta los siguientes lemas:

Lema 3.4 (Basado en el Lema 7.5 de [3]).

Sea $\|\cdot\|$ una norma en \mathbb{R}^d y sea $\varepsilon \in (0, 1)$. Si denotamos $S := \{x \in \mathbb{R}^d : \|x\| = 1\}$, entonces existe un conjunto $Q \subseteq S$ tal que $\text{card}(Q) \leq (3/\varepsilon)^d$ y tal que para cualquier $x \in S$ existe un punto $q \in Q$ cumpliendo $\|x - q\| \leq \varepsilon$.

Demostración. Construiremos Q de manera recurrente. Para empezar tomamos un $q_1 \in S$. Vamos añadiendo puntos a Q de forma que en el paso i añadimos un punto $q_i \in S$ que cumple $\|q_i - q_j\| > \varepsilon$ para todo $j < i$. Así continuamos hasta que no podamos añadir más puntos. De esta manera, para cualquier $x \in S$ existirá un punto $q \in Q$ tal que $\|x - q\| \leq \varepsilon$. Pues si no ocurriera esto, entonces podríamos añadir ese x a Q .

Ahora falta acotar $\text{card}(Q)$. Observamos que las bolas de radio $\varepsilon/2$ centradas en los distintos puntos de Q , $B(q, \varepsilon/2)$, son disjuntas y que además su unión está contenida en la bola centrada en el origen de radio $1 + \varepsilon/2$, $B(0, 1 + \varepsilon/2)$. Entonces,

$$\bigcup_{q \in Q} B(q, \varepsilon/2) \subset B(0, 1 + \varepsilon/2),$$

y tomando volúmenes

$$\text{card}(Q) \text{vol}(B(0, \varepsilon/2)) = \text{vol}\left(\bigcup_{q \in Q} B(q, \varepsilon/2)\right) \leq \text{vol}(B(0, 1 + \varepsilon/2)).$$

Por tanto,

$$\text{card}(Q) \leq \frac{\text{vol}(B(0, 1 + \varepsilon/2))}{\text{vol}(B(0, \varepsilon/2))} = \frac{(1 + \varepsilon/2)^d}{(\varepsilon/2)^d} = \left(\frac{2 + \varepsilon}{\varepsilon}\right)^d \underset{0 < \varepsilon < 1}{\leq} \left(\frac{3}{\varepsilon}\right)^d.$$

□

Dado un conjunto de índices T en $\{1, \dots, N\}$ denotamos por X_T el subespacio de vectores de \mathbb{R}^N cuyas entradas fuera de T son 0.

Lema 3.5. Sea $\Phi(\omega)$, $\omega \in \Omega$, una matriz aleatoria de tamaño $m \times N$ obtenida de acuerdo a una distribución que cumple la propiedad de Concentración de Medida. Entonces, dado un entero positivo $s < m$ y $0 < \delta < 1$, para cualquier conjunto de índices $T \subset \{1, \dots, N\}$ con $\text{card}(T) \leq s$ tenemos

$$(1 - \delta)\|x\|_2 \leq \|\Phi(\omega)x\|_2 \leq (1 + \delta)\|x\|_2, \quad \forall x \in X_T, \quad (3.15)$$

con probabilidad $\geq 1 - 2\left(\frac{12}{\delta}\right)^s e^{-mc_o(\delta/2)}$.

Demostración. Podemos centrarnos primero en los $x \in X_T$ con $\|x\|_2 = 1$.

Escogemos una colección finita de puntos Q_T tal que $Q_T \subset X_T$, $\|q\|_2 = 1$ para todo $q \in Q_T$ y tal que para todo $x \in X_T$ con $\|x\|_2 = 1$ tenemos $\|x - q\|_2 \leq \delta/4$, para algún $q \in Q_T$. Por el lema anterior, con $\varepsilon = \delta/4$, sabemos que Q_T existe y que además, lo podemos elegir con

$$|Q_T| = \text{card}(Q_T) \leq \left(\frac{3}{\delta/4}\right)^s = \left(\frac{12}{\delta}\right)^s.$$

Si ahora llamamos

$$\Omega(q) = \left\{ \omega \in \Omega : \left| \|\Phi(\omega)q\|_2^2 - \|q\|_2^2 \right| > \frac{\delta}{2} \|q\|_2^2 \right\},$$

como se cumple la propiedad de Concentración de Medida, tomando $\varepsilon = \delta/2$, para cada $q \in Q_T$, se tiene

$$\mathbb{P}(\Omega(q)) = \mathbb{P} \left(\left| \|\Phi(\omega)q\|_2^2 - \|q\|_2^2 \right| > \frac{\delta}{2} \|q\|_2^2 \right) \leq 2e^{-mc_0(\delta/2)}.$$

Usando las propiedades de la unión y su probabilidad, calculamos lo siguiente:

$$\begin{aligned} \mathbb{P} \left(\left| \|\Phi(\omega)q\|_2^2 - \|q\|_2^2 \right| \leq \frac{\delta}{2} \|q\|_2^2, \forall q \in Q_T \right) &= \mathbb{P} \left(\left(\bigcup_{q \in Q_T} \Omega(q) \right)^c \right) = 1 - \mathbb{P} \left(\bigcup_{q \in Q_T} \Omega(q) \right) \geq \\ &\geq 1 - \sum_{q \in Q_T} \mathbb{P}(\Omega(q)) \geq 1 - |Q_T| 2e^{-mc_0(\delta/2)} \geq 1 - 2 \left(\frac{12}{\delta} \right)^s e^{-mc_0(\delta/2)}. \end{aligned}$$

Tenemos entonces, con probabilidad $\geq 1 - 2 \left(\frac{12}{\delta} \right)^s e^{-mc_0(\delta/2)}$,

$$\begin{aligned} \left| \|\Phi(\omega)q\|_2^2 - \|q\|_2^2 \right| &\leq \frac{\delta}{2} \|q\|_2^2, \forall q \in Q_T. \\ -\frac{\delta}{2} \|q\|_2^2 &\leq \|\Phi(\omega)q\|_2^2 - \|q\|_2^2 \leq \frac{\delta}{2} \|q\|_2^2, \forall q \in Q_T. \\ \left(1 - \frac{\delta}{2} \right) \|q\|_2^2 &\leq \|\Phi(\omega)q\|_2^2 \leq \left(1 + \frac{\delta}{2} \right) \|q\|_2^2, \forall q \in Q_T. \\ \sqrt{1 - \frac{\delta}{2}} \|q\|_2 &\leq \|\Phi(\omega)q\|_2 \leq \sqrt{1 + \frac{\delta}{2}} \|q\|_2, \forall q \in Q_T. \end{aligned}$$

Como $\delta \in (0, 1)$ se cumple

$$1 - \delta/2 \leq \sqrt{1 - \delta/2} \quad \text{y} \quad \sqrt{1 + \delta/2} \leq 1 + \delta/2.$$

Y nos queda

$$\left(1 - \frac{\delta}{2} \right) \|q\|_2 \leq \|\Phi(\omega)q\|_2 \leq \left(1 + \frac{\delta}{2} \right) \|q\|_2, \forall q \in Q_T. \quad (3.16)$$

Consideramos $A = A(\omega) \geq 0$ el número más pequeño tal que $\|\Phi(\omega)x\|_2 \leq (1 + A)\|x\|_2$, $\forall x \in X_T$. Vamos a ver que $A \leq \delta$. Recordamos que para cualquier $x \in X_T$ con $\|x\|_2 = 1$, existe un $q \in Q_T$ tal que $\|x - q\|_2 \leq \delta/4$. Así,

$$\begin{aligned} \|\Phi(\omega)x\|_2 &\leq \|\Phi(\omega)q\|_2 + \|\Phi(\omega)(x - q)\|_2 \leq \leftarrow (3.16) \\ &\leq \left(1 + \frac{\delta}{2} \right) \|q\|_2 + (1 + A)\|x - q\|_2 \leq 1 + \frac{\delta}{2} + (1 + A)\frac{\delta}{4}. \end{aligned}$$

Por otro lado, para $x \in X_T$ con $\|x\|_2 = 1$ también se cumple $\|\Phi(\omega)x\|_2 \leq 1 + A$ y como A es el número más pequeño que lo cumple, entonces

$$A \leq \frac{\delta}{2} + (1 + A)\frac{\delta}{4} \quad \implies \quad A \leq \frac{\frac{3}{4}\delta}{1 - \frac{\delta}{4}} \leq \frac{\frac{3}{4}\delta}{1 - \frac{1}{4}} = \delta.$$

Tenemos ya la cota superior. Para la inferior usaremos que

$$\|\Phi(\omega)q\|_2 - \|\Phi(\omega)(q-x)\|_2 \leq \|\Phi(\omega)x\|_2$$

y, entonces,

$$\|\Phi(\omega)x\|_2 \underset{(3.16)}{\geq} \left(1 - \frac{\delta}{2}\right) - (1 + \delta)\frac{\delta}{4} \geq 1 - \frac{\delta}{2} - \frac{\delta}{4} - \frac{\delta^2}{4} \underset{\delta \in (0,1)}{\geq} 1 - \frac{\delta}{2} - \frac{\delta}{4} - \frac{\delta}{4} = 1 - \delta.$$

En definitiva, con probabilidad $\geq 1 - 2\left(\frac{12}{\delta}\right)^s e^{-mc_0(\delta/2)}$ se cumple

$$1 - \delta \leq \|\Phi(\omega)x\|_2 \leq 1 + \delta, \quad \forall x \in X_T \text{ con } \|x\|_2 = 1.$$

Para el resto de $x \in X_T$ bastará hacer

$$\|\Phi(\omega)x\|_2 = \left\| \Phi(\omega) \frac{x}{\|x\|_2} \|x\|_2 \right\|_2 = \|x\|_2 \left\| \Phi(\omega) \frac{x}{\|x\|_2} \right\|_2.$$

Por tanto, con probabilidad $\geq 1 - 2\left(\frac{12}{\delta}\right)^s e^{-mc_0(\delta/2)}$ se cumple

$$(1 - \delta)\|x\|_2 \leq \|\Phi(\omega)x\|_2 \leq (1 + \delta)\|x\|_2, \quad \forall x \in X_T.$$

□

Tanto en este lema como en el siguiente resultado, el entero s representa cómo de *sparse* son las señales que tratamos. Como decíamos al principio, todo este desarrollo se hace partiendo de que sabemos que las señales que queremos reconstruir son suficientemente *sparse*, o sea, s es suficientemente pequeño. Podemos entonces suponer que $s < m < N$ es lo bastante pequeño para que $\log(N/s) > 1$, es decir, $s < N/e$.

Bajo estas condiciones tenemos ya el resultado que queríamos demostrar.

Teorema 3.6. *Sea $\Phi(\omega)$, $\omega \in \Omega$, una matriz aleatoria de tamaño $m \times N$ obtenida de acuerdo a una distribución que cumple la propiedad de Concentración de Medida. Dado $0 < \delta < 1$, existen constantes $c_1, c_2 > 0$, que dependen solo de δ , tales que para cualquier $s \leq \frac{c_1 m}{\log(N/s)}$ con probabilidad $\geq 1 - 2e^{-c_2 m}$, $\Phi(\omega)$ cumple*

$$(1 - \delta)\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq (1 + \delta)\|x\|_2^2, \quad \forall x \in \Sigma_s.$$

Demostración. Tomamos $\delta' = \delta/3 \in (0, 1)$. Denotaremos $[N] := \{1, \dots, N\}$. Sabemos del Lema 3.5 que para cada $T \subset [N]$ con $|T| = \text{card}(T) = s$, y cada uno de los espacios s -dimensionales, X_T , la matriz $\Phi(\omega)$ no cumplirá

$$(1 - \delta')\|x\|_2 \leq \|\Phi(\omega)\|_2 \leq (1 + \delta')\|x\|_2, \quad \forall x \in X_T,$$

con probabilidad $\leq 2\left(\frac{12}{\delta'}\right)^s e^{-mc_0(\delta'/2)}$.

Hay $\binom{N}{s}$ de esos subespacios y se tiene $\binom{N}{s} \leq (eN/s)^s$. En efecto,

$$\begin{aligned} \binom{N}{s} \left(\frac{s}{N}\right)^s &= \frac{N!}{s!(N-s)!} \left(\frac{s}{N}\right)^s = \frac{N(N-1)\cdots(N-s+1)}{N^s} \frac{s^s}{s!} \leq \frac{s^s}{s!} \leq \\ &\leq 1 + s + \frac{s^2}{2!} + \cdots + \frac{s^s}{s!} + \cdots = e^s. \end{aligned}$$

Definimos el conjunto

$$\Omega(T) := \{\omega \in \Omega : (1 - \delta')\|x\|_2 \leq \|\Phi(\omega)\|_2 \leq (1 + \delta')\|x\|_2, \forall x \in X_T\}.$$

Entonces tendremos:

$$\begin{aligned} \mathbb{P}\left((1 - \delta')\|x\|_2 \leq \|\Phi(\omega)\|_2 \leq (1 + \delta')\|x\|_2, \forall x \in \Sigma_s\right) &= \mathbb{P}\left(\bigcap_{\substack{T \subset [N] \\ |T|=s}} \Omega(T)\right) = \\ &= 1 - \mathbb{P}\left(\bigcup_{\substack{T \subset [N] \\ |T|=s}} \Omega(T)^c\right) \geq 1 - \sum_{\substack{T \subset [N] \\ |T|=s}} \mathbb{P}(\Omega(T)^c) \geq \\ &\geq 1 - 2 \binom{N}{s} \left(\frac{12}{\delta'}\right)^s e^{-mc_0(\delta'/2)} \geq 1 - 2(eN/s)^s \left(\frac{12}{\delta'}\right)^s e^{-mc_0(\delta'/2)} = \\ &= 1 - 2 \exp\left(-mc_0(\delta'/2) + s \left[\log(eN/s) + \log(12/\delta')\right]\right) = \\ &= 1 - 2 \exp\left(-mc_0(\delta'/2) + s \left[1 + \log(N/s) + \log(12/\delta')\right]\right). \end{aligned}$$

Si tomamos $c_1 := \frac{1}{2} \cdot \frac{c_0(\delta'/2)}{2 + \log(12/\delta')} > 0$, cuando $s \leq \frac{c_1 m}{\log(N/s)}$ se tiene

$$\begin{aligned} \exp\left(-mc_0(\delta'/2) + s \left[1 + \log(N/s) + \log(12/\delta')\right]\right) &\leq \\ &\leq \exp\left(-mc_0(\delta'/2) + \frac{c_1 m}{\log(N/s)} \left[1 + \log(N/s) + \log(12/\delta')\right]\right) = \\ &= \exp\left(-m \left[c_0(\delta'/2) - c_1 \left(1 + \frac{1 + \log(12/\delta')}{\log(N/s)}\right)\right]\right) \leq \leftarrow{(\log(N/s) \geq 1)} \\ &\leq \exp\left(-m \left[c_0(\delta'/2) - c_1 (2 + \log(12/\delta'))\right]\right) = \\ &= \exp\left(-m \frac{c_0(\delta'/2)}{2}\right) \leq \exp(-mc_2), \end{aligned}$$

siempre que $c_2 \geq \frac{c_0(\delta'/2)}{2} > 0$.

Así, tendremos que con probabilidad $\geq 1 - 2e^{-mc_2}$ se cumple

$$(1 - \delta')\|x\|_2 \leq \|\Phi(\omega)\|_2 \leq (1 + \delta')\|x\|_2, \forall x \in \Sigma_s.$$

Veamos que esto implica que se cumplen las desigualdades al cuadrado para el δ dado inicialmente. Si elevamos todo al cuadrado, tendremos

$$(1 - \delta')^2 \|x\|_2^2 \leq \|\Phi(\omega)x\|_2^2 \leq (1 + \delta')^2 \|x\|_2^2, \forall x \in \Sigma_s.$$

Por un lado,

$$\|\Phi(\omega)x\|_2^2 \leq (1 + \delta')^2 \|x\|_2^2 = (1 + 2\delta' + (\delta')^2) \|x\|_2^2 \stackrel{\delta' \in (0,1)}{\leq} (1 + 3\delta') \|x\|_2^2 \stackrel{\delta' = \delta/3}{\leq} (1 + \delta) \|x\|_2^2.$$

Por otro,

$$\|\Phi(\omega)x\|_2^2 \geq (1 - \delta')^2 \|x\|_2^2 = (1 - 2\delta' + (\delta')^2) \|x\|_2^2 \geq (1 - 3\delta') \|x\|_2^2 = (1 - \delta) \|x\|_2^2.$$

En definitiva, para $0 < \delta < 1$ y $s \leq \frac{c_1 m}{\log(N/s)}$ con probabilidad $\geq 1 - 2e^{-mc_2}$ se cumple

$$(1 - \delta) \|x\|_2^2 \leq \|\Phi(\omega)x\|_2^2 \leq (1 + \delta) \|x\|_2^2, \forall x \in \Sigma_s.$$

□

Corolario 3.7. *Por la definición de la propiedad RIP basta aplicar el teorema anterior con un $\delta < 1$ y tendremos probado que la propiedad de Concentración de Medida implica RIP con probabilidad alta. Si, además, pedimos $\delta < \sqrt{2} - 1$ con $2s$ en vez de s se cumplirán los teoremas de reconstrucción de la Minimización ℓ_1 ; ver Teoremas 3.1 y 3.2.*

3. Número de mediciones

La condición $s \leq \frac{c_1 m}{\log(N/s)}$ del teorema anterior se puede interpretar de distinta manera según nuestra situación.

Podría ser que realizar mediciones para obtener más información sea muy costoso y, por tanto, solo dispongamos de un número limitado de ellas, es decir, que m sea fijo. Entonces utilizando esta relación podríamos obtener el máximo s para el cual somos capaces de recuperar de manera exacta las señales s -sparse.

O, por el contrario, podríamos conocer ya de antemano cómo de *sparse* son las señales que queremos recuperar, o sea, conocemos s , y utilizar la relación para obtener el número mínimo de mediciones necesarias para reconstruir las señales.

$$s \leq \frac{c_1 m}{\log(N/s)} \implies c_1^{-1} s \log(N/s) \leq m. \quad (*)$$

De hecho, vamos a probar que el número de mediciones necesarias para que exista una matriz $m \times N$ que permita reconstruir señales s -sparse usando la Minimización ℓ_1 debe necesariamente cumplir $m \geq c \cdot s \log(N/s)$, para algún $c > 0$. Con lo que el valor obtenido en (*) será óptimo (salvo quizá por la constante c_1). Este resultado lo podemos encontrar en el Teorema 10.11 del libro [11].

Teorema 3.8 (Cota para el número de mediciones m).

Dada una matriz $A \in \mathbb{R}^{m \times N}$, si todo vector $2s$ -sparse $x \in \mathbb{R}^N$ minimiza $\|z\|_1$ sujeto a que $Az = Ax$, entonces

$$m \geq c_1 s \log\left(\frac{N}{c_2 s}\right), \quad \text{donde } c_1 = \frac{1}{\log(9)} \text{ y } c_2 = 4. \quad (3.17)$$

Para probarlo necesitaremos el siguiente lema combinatorio:

Lema 3.9. *Dados enteros positivos $s < N$, existen $n \geq \left(\frac{N}{4s}\right)^{s/2}$ subconjuntos S_1, \dots, S_n de $[N] = \{1, \dots, N\}$ tales que cada S_i tiene cardinal s y $|S_i \cap S_j| < s/2$ para $i \neq j$.*

Demostración. Si $s > N/4$ entonces $\left(\frac{N}{4s}\right)^{s/2} < 1$ y bastaría con coger un único subconjunto S de $[N]$ con cardinal s .

Suponemos entonces que $s \leq N/4$. Denotamos por $B := \{S \subseteq [N] : |S| = s\}$. Elegimos un $S_1 \in B$ y denotamos $A_1 := \{S \in B : |S_1 \cap S| \geq s/2\}$. Se tiene

$$\begin{aligned} |A_1| &= \sum_{k=\lceil s/2 \rceil}^s \binom{s}{k} \binom{N-s}{s-k} \leq \max_{\lceil s/2 \rceil \leq k \leq s} \binom{N-s}{s-k} \cdot \sum_{k=\lceil s/2 \rceil}^s \binom{s}{k} \leq \\ &\leq \max_{\lceil s/2 \rceil \leq k \leq s} \binom{N-s}{s-k} \cdot \sum_{k=0}^s \binom{s}{k} = 2^s \max_{\lceil s/2 \rceil \leq k \leq s} \binom{N-s}{s-k} = \\ &= 2^s \max_{0 \leq t \leq \lfloor s/2 \rfloor} \binom{N-s}{t}. \end{aligned}$$

Sabemos que los números combinatorios cumplen $\binom{m}{l} \leq \binom{m}{l+1}$, si $l < \lfloor m/2 \rfloor$. Y, por tanto, $|A_1| \leq 2^s \binom{N-s}{\lfloor s/2 \rfloor}$. Observamos que todo $S \in B \setminus A_1$ cumple $|S| = s$ y $|S_1 \cap S| < s/2$. Podemos entonces, suponiendo que $B \setminus A_1 \neq \emptyset$, tomar $S_2 \in B \setminus A_1$.

Como hacíamos antes consideramos $A_2 := \{S \in B \setminus A_1 : |S_2 \cap S| \geq s/2\}$ que también cumplirá $|A_2| \leq 2^s \binom{N-s}{\lfloor s/2 \rfloor}$. Cualquier conjunto $S \in B \setminus (A_1 \cup A_2)$ cumplirá $|S| = s$, $|S_1 \cap S| < s/2$ y $|S_2 \cap S| < s/2$. Entonces, si todavía $B \setminus (A_1 \cup A_2) \neq \emptyset$, podemos tomar $S_3 \in B \setminus (A_1 \cup A_2)$.

Seguiríamos repitiendo el proceso seleccionando conjuntos S_1, \dots, S_n y las respectivas familias A_1, \dots, A_n hasta que $B \setminus (A_1 \cup \dots \cup A_n) = \emptyset$. De esta manera, tendríamos n subconjuntos de $[N]$ todos con cardinal s y que además cumplen $|S_i \cap S_j| < s/2$ para $i \neq j$. Por construcción, todos los A_i cumplen

$$|A_i| \leq 2^s \binom{N-s}{\lfloor s/2 \rfloor}.$$

Nos falta acotar el número de subconjuntos, n . Como $B \setminus (A_1 \cup \dots \cup A_n) = \emptyset$, entonces

$$\binom{N}{s} = |B| = \left| \bigcup_{i=1}^n A_i \right| \leq \sum_{i=1}^n |A_i| \leq n \max_{1 \leq i \leq n} |A_i| \leq n 2^s \binom{N-s}{\lfloor s/2 \rfloor}.$$

Y así,

$$\begin{aligned} n &\geq \frac{\binom{N}{s}}{2^s \binom{N-s}{\lfloor s/2 \rfloor}} = \frac{1}{2^s} \frac{N!}{s!(N-s)!} \frac{\lfloor s/2 \rfloor! (N-s-\lfloor s/2 \rfloor)!}{(N-s)!} = \\ &= \frac{1}{2^s} \frac{N(N-1)\dots(N-s+1)}{s(s-1)\dots(s-\lfloor s/2 \rfloor+1)(N-s)(N-s-1)\dots(N-s-\lfloor s/2 \rfloor+1)} \stackrel{\lfloor s/2 \rfloor \leq \lfloor s/2 \rfloor}{\geq} \\ &\geq \frac{1}{2^s} \frac{N(N-1)\dots(N-\lfloor s/2 \rfloor+1)}{s(s-1)\dots(s-\lfloor s/2 \rfloor+1)} \frac{(N-\lfloor s/2 \rfloor)\dots(N-\overbrace{\lfloor s/2 \rfloor}^{-s}+1)}{(N-s)(N-s-1)\dots(N-s-\lfloor s/2 \rfloor+1)} \geq \\ &\geq \frac{1}{2^s} \frac{N(N-1)\dots(N-\lfloor s/2 \rfloor+1)}{s(s-1)\dots(s-\lfloor s/2 \rfloor+1)} \geq \frac{1}{2^s} \left(\frac{N}{s}\right)^{\lfloor s/2 \rfloor} \geq \left(\frac{N}{4s}\right)^{s/2}. \end{aligned}$$

Donde en la última línea hemos utilizado la desigualdad,

$$\frac{N-j}{s-j} \geq \frac{N}{s}, \quad \forall j = 0, \dots, \lfloor s/2 \rfloor - 1,$$

que se justifica usando que la función $f(t) = \frac{N-t}{s-t} = 1 + \frac{N-s}{s-t}$ es claramente creciente si $t < s$. □

También nos hará falta el siguiente resultado de análisis funcional que podemos encontrar como la Proposición 1.1.11 en el libro [9].

Proposición 3.10. *Sea X un espacio normado. Si Y es un subespacio vectorial cerrado del espacio normado X , entonces el espacio vectorial cociente X/Y es un espacio normado para la norma cociente dada por*

$$|x + Y| = \inf\{\|x + y\| : y \in Y\}.$$

Con todo esto ya estamos en condiciones de probar el Teorema 3.8.

Dem. del Teorema 3.8. Recordamos que estamos trabajando con el espacio normado ℓ_1^N y dada una matriz $A \in \mathbb{R}^{m \times N}$ sabemos por álgebra lineal que su núcleo,

$$\ker A = \{x \in \mathbb{R}^N : Ax = 0\},$$

es un subespacio vectorial cerrado de ℓ_1^N . Así, por el resultado de análisis funcional, podemos considerar el espacio cociente $X := \ell_1^N / \ker A = \{[x] := x + \ker A, x \in \mathbb{R}^N\}$ normado con

$$\|[x]\| := \inf_{v \in \ker A} \|x - v\|_1, \quad x \in \mathbb{R}^N.$$

Dado un vector $2s$ -sparse $x \in \mathbb{R}^N$, observamos que todo vector $z = x - v$ con $v \in \ker A$ satisface $Az = Ax$. Como por hipótesis, los vectores $x \in \Sigma_{2s}$ minimizan la norma ℓ_1 en el conjunto $x + \ker A$, tendremos $\|[x]\| = \|x\|_1$.

Sean S_1, \dots, S_n los conjuntos introducidos en el lema anterior. Como tienen cardinal s , podemos definir vectores s -sparse $x^1, \dots, x^n \in \mathbb{R}^N$ con norma ℓ_1 igual a 1 haciendo

$$x^i = (1/s)\mathbf{1}_{S_i}, \text{ entendienddo } \mathbf{1} = (1, \dots, 1) \in \mathbb{R}^N.$$

Para $1 \leq i \neq j \leq n$ tenemos que $x^i - x^j$ es $2s$ -sparse y entonces sabemos que

$$\|[x^i] - [x^j]\| = \|[x^i - x^j]\| = \|x^i - x^j\|_1.$$

Además, por la construcción de estos vectores, se tiene

$$|x_k^i - x_k^j| = \begin{cases} 1/s & \text{si } k \in (S_i \cup S_j) \setminus (S_i \cap S_j). \\ 0 & \text{en otro caso.} \end{cases}$$

Calculando el cardinal obtenemos

$$\begin{aligned} |(S_i \cup S_j) \setminus (S_i \cap S_j)| &= |S_i \cup S_j| - |S_i \cap S_j| = |S_i| + |S_j| - |S_i \cap S_j| - |S_i \cap S_j| = \\ &= \underbrace{|S_i|}_s + \underbrace{|S_j|}_s - 2 \underbrace{|S_i \cap S_j|}_{\wedge_{s/2}} > 2s - 2 \frac{s}{2} = 2s - s = s. \end{aligned}$$

Con lo que, $\|x^i - x^j\|_1 > 1$, y entonces, $\|[x^i] - [x^j]\| > 1$ para todo $1 \leq i \neq j \leq n$.

Esto muestra que las bolas de radio $1/2$ centradas en los puntos $[x^i]$ son disjuntas y, además, al tener al tener los x^i norma 1, estarán contenidas en la bola de radio $1 + 1/2$ centrada en el origen,

$$\bigcup_{i=1}^n B([x^i], 1/2) \subseteq B(0, 3/2).$$

Si comparamos volúmenes,

$$n \cdot \text{vol}(B(0, 1/2)) = \text{vol} \left(\bigcup_{i=1}^n B([x^i], 1/2) \right) \leq \text{vol}(B(0, 3/2)).$$

Como estas bolas son subconjuntos del espacio cociente X , tendremos

$$n \leq \frac{\text{vol}(B(0, 3/2))}{\text{vol}(B(0, 1/2))} = \frac{(3/2)^r}{(1/2)^r} = 3^r \leq 3^m, \quad \text{donde } r = \dim(X) = \text{rank}(A) \leq m.$$

El lema anterior también nos decía que $\left(\frac{N}{4s}\right)^{s/2} \leq n \leq 3^m$ y tomando logaritmos obtenemos:

$$\begin{aligned} \frac{s}{2} \log\left(\frac{N}{4s}\right) &= \log\left(\left(\frac{N}{4s}\right)^{s/2}\right) \leq \log(3^m) = m \log(3) \\ \implies s \log\left(\frac{N}{4s}\right) &\leq 2m \log(3) = m \log(9) \\ \implies \frac{s}{\log(9)} \log\left(\frac{N}{4s}\right) &\leq m \\ \implies c_1 s \log\left(\frac{N}{c_2 s}\right) &\leq m. \end{aligned}$$

□

Cambiando $2s$ por s , y salvo por constantes, este teorema nos indica una cota inferior para el número de mediciones necesarias para reconstruir de manera exacta las señales s -sparse con el algoritmo ℓ_1 . Decíamos que el método de las matrices aleatorias era óptimo pues su número de mediciones es del mismo orden que esta cota.

Capítulo 4

Distribución Subgaussiana

A lo largo de este capítulo estudiaremos las variables aleatorias Subgaussianas. Veremos su definición, cómo caracterizarlas y algunas de sus propiedades más relevantes. El principal resultado que probaremos es que las matrices generadas a partir de un tipo concreto de variable Subgaussiana cumplen la Concentración de Medida. Usando esta propiedad seremos capaces de probar el célebre Lema de Johnson-Lindenstrauss. También proporcionaremos, al final del capítulo, algunos ejemplos concretos de distribuciones Subgaussianas.

Para este capítulo usaremos varios conceptos de Teoría de la Probabilidad que podemos encontrar resumidos en el Apéndice C.

1. Definición y momentos

Definición 4.1 (Subgaussiana).

Una variable aleatoria X diremos que es Subgaussiana si existe una constante $c > 0$ tal que

$$M_X(t) = \mathbb{E}[e^{tX}] \leq e^{\frac{c^2 t^2}{2}}, \quad \forall t \in \mathbb{R}.$$

La denotaremos $X \sim \text{Sub}(c^2)$.

Observación. Recibe el nombre de Subgaussiana pues $e^{\frac{c^2 t^2}{2}}$ coincide con la función generatriz de momentos de una variable Normal de media 0 y varianza c^2 . Lo veremos en la sección de ejemplos del final del capítulo, donde probaremos que, si $X \sim \text{Normal}(0, \sigma^2)$ entonces $X \sim \text{Sub}(\sigma^2)$.

Lo primero que vamos a probar son algunas propiedades sobre los momentos de las distribuciones Subgaussianas. Para ello necesitaremos algunos lemas.

Lema 4.1. Si $X \sim \text{Sub}(c^2)$ y $k \in \mathbb{N}$, entonces $\exists A_k = 2(k/e)^{k/2}$ tal que

$$\mathbb{E}[|X|^k] \leq A_k c^k.$$

Demostración. Consideramos $g_k(t) = t^k e^{-t}$, $t \geq 0$ que alcanza su máximo en $t = k$. En efecto,

$$g'_k(t) = kt^{k-1}e^{-t} + t^k e^{-t}(-1) = t^{k-1}e^{-t}(k-t) = 0 \iff t = k.$$

Luego,

$$t^k e^{-t} \leq k^k e^{-k} = (k/e)^k, \quad \forall t \geq 0.$$

Para $\lambda > 0$ tendremos

$$\begin{aligned} \mathbb{E}[|X|^k] &= \mathbb{E}[|X|^k e^{-\lambda|X|} e^{\lambda|X|}] = \frac{1}{\lambda^k} \mathbb{E}[(\lambda|X|)^k e^{-\lambda|X|} e^{\lambda|X|}] \leq \frac{1}{\lambda^k} \mathbb{E}[(k/e)^k e^{\lambda|X|}] = \\ &= \left(\frac{k}{\lambda e}\right)^k \mathbb{E}[e^{\lambda|X|}] \leq \left(\frac{k}{\lambda e}\right)^k \mathbb{E}[e^{\lambda X} + e^{-\lambda X}] \leq \left(\frac{k}{\lambda e}\right)^k 2e^{\frac{c^2 \lambda^2}{2}}. \end{aligned}$$

Minimizamos respecto de $\lambda > 0$. Consideramos $g(\lambda) = \left(\frac{k}{\lambda e}\right)^k 2e^{\frac{c^2\lambda^2}{2}}$. Entonces,

$$g'(\lambda) = \frac{-k}{\lambda^{k+1}} \left(\frac{k}{e}\right)^k 2e^{\frac{c^2\lambda^2}{2}} + c^2\lambda \left(\frac{k}{\lambda e}\right)^k 2e^{\frac{c^2\lambda^2}{2}} = \frac{1}{\lambda^{k+1}} \left(\frac{k}{e}\right)^k 2e^{\frac{c^2\lambda^2}{2}} [-k + c^2\lambda^2] = 0$$

$$\iff \lambda^2 = \frac{k}{c^2} \iff \lambda = \frac{\sqrt{k}}{c} > 0.$$

Por tanto, el mínimo se alcanza en $\lambda = \frac{\sqrt{k}}{c}$ y nos queda

$$\mathbb{E}[|X|^k] \leq \left(\frac{k}{\frac{\sqrt{k}}{c}e}\right)^k 2e^{\frac{c^2\frac{k}{c^2}}{2}} = \left(\frac{\sqrt{k}c}{e}\right)^k 2e^{\frac{k}{2}} = \underbrace{2\left(\frac{k}{e}\right)^{k/2}}_{A_k} c^k = A_k c^k.$$

□

Observación. Con las constantes $A_k = 2(k/e)^{k/2}$, la serie de potencias

$$S(t) := 1 + \sum_{k=1}^{\infty} A_k \frac{t^k}{k!}$$

tiene radio de convergencia $R = \infty$.

En efecto, usando el Criterio de la raíz y la Fórmula de Stirling, $k! \simeq (k/e)^k \sqrt{2\pi k}$.

$$\left(\frac{A_k}{k!}\right)^{1/k} \sim \frac{2^{\frac{1}{k}} \sqrt{k/e}}{(2\pi k)^{\frac{1}{2k}} (k/e)} = \frac{2^{\frac{1}{k}}}{(2\pi k)^{\frac{1}{2k}}} \frac{1}{\sqrt{k/e}} \xrightarrow{k \rightarrow \infty} 0 \implies R = \infty.$$

Esta observación junto con el lema permiten justificar la validez de

$$\mathbb{E}[e^{tX}] = 1 + \sum_{k=1}^{\infty} \frac{\mathbb{E}[(tX)^k]}{k!}, \quad \forall t \in \mathbb{R},$$

ya que

$$\left| \sum_{k=1}^{\infty} \frac{\mathbb{E}[(tX)^k]}{k!} \right| \leq \sum_{k=1}^{\infty} \frac{|t|^k}{k!} \mathbb{E}[|X|^k] \leq \sum_{k=1}^{\infty} \frac{A_k}{k!} |t|^k c^k = S(|t|c) < \infty,$$

pues la serie $S(\cdot)$ es absolutamente convergente.

De hecho, podemos probar el siguiente lema:

Lema 4.2. Sea $X \sim \text{Sub}(c^2)$. Entonces, para cada $n \in \mathbb{N}$, se tiene

$$\mathbb{E}[e^{tX}] = 1 + \sum_{k=1}^n \frac{\mathbb{E}[(tX)^k]}{k!} + \mathcal{O}(|t|^{n+1}), \quad \text{si } |t| \leq 1.$$

Demostración. Notar que si $|t| \leq 1$,

$$\sum_{k \geq n+1} \frac{A_k |t|^k c^k}{k!} \leq |t|^{n+1} \sum_{k \geq n+1} \frac{A_k c^k}{k!} \leq |t|^{n+1} S(c).$$

Por tanto, para $|t| \leq 1$,

$$\begin{aligned} \left| \mathbb{E} \left[e^{tX} - \sum_{k=0}^n \frac{(tX)^k}{k!} \right] \right| &\leq \mathbb{E} \left[\left| e^{tX} - \sum_{k=0}^n \frac{(tX)^k}{k!} \right| \right] = \mathbb{E} \left[\left| \sum_{k \geq n+1} \frac{(tX)^k}{k!} \right| \right] \leq \\ &\leq \mathbb{E} \left[\sum_{k \geq n+1} \frac{|tX|^k}{k!} \right] = \sum_{k \geq n+1} \frac{|t|^k}{k!} \mathbb{E} [|X|^k] \leq \sum_{k \geq n+1} \frac{|t|^k}{k!} A_k c^k \leq |t|^{n+1} S(c). \end{aligned}$$

Y llegamos a lo que queríamos,

$$\mathbb{E}[e^{tX}] = 1 + \sum_{k=1}^n \frac{\mathbb{E}[(tX)^k]}{k!} + \mathcal{O}(|t|^{n+1}), \quad \text{si } |t| \leq 1.$$

□

Usando este lema podemos probar la siguiente proposición, que nos indica cómo son los momentos de primer y segundo orden.

Proposición 4.3. Si $X \sim \text{Sub}(c^2)$, entonces $\mathbb{E}[X] = 0$ y $\mathbb{E}[X^2] \leq c^2$.

Demostración. Por definición de Subgaussiana sabemos

$$\mathbb{E}[e^{tX}] \leq e^{\frac{c^2 t^2}{2}} = \sum_{n \geq 0} \left(\frac{c^2 t^2}{2} \right)^n \frac{1}{n!}.$$

Gracias al lema podemos desarrollar el lado izquierdo también como serie y, si consideramos solo los términos hasta grado 2, tendremos

$$1 + t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}[X^2] + \mathcal{O}(|t|^3) \leq 1 + \frac{c^2 t^2}{2} + \mathcal{O}(|t|^4).$$

Vamos a empezar calculando $\mathbb{E}[X]$, así que, de la expresión anterior pasamos a

$$t\mathbb{E}[X] \leq \frac{t^2}{2} (c^2 - \mathbb{E}[X^2]) + \mathcal{O}(|t|^3) = \mathcal{O}(|t|^2).$$

- Para $t > 0$, dividiendo por t , $\mathbb{E}[X] \leq \mathcal{O}(|t|)$ y tomando $t \rightarrow 0^+$ tenemos $\mathbb{E}[X] \leq 0$.
- Para $t < 0$, dividiendo por t , $\mathbb{E}[X] \geq \mathcal{O}(|t|)$ y tomando $t \rightarrow 0^-$ tenemos $\mathbb{E}[X] \geq 0$.

Por tanto, $\mathbb{E}[X] = 0$ y nos quedará $\frac{t^2}{2}\mathbb{E}[X^2] \leq \frac{c^2 t^2}{2} + \mathcal{O}(|t|^3)$.

Para $t \neq 0$, dividiendo por $\frac{t^2}{2}$, $\mathbb{E}[X^2] \leq c^2 + \mathcal{O}(|t|)$ y tomando $t \rightarrow 0$, entonces

$$\mathbb{E}[X^2] \leq c^2.$$

□

Hemos visto que si $X \sim \text{Sub}(c^2)$, entonces $\mathbb{E}[X^2] \leq c^2$. No obstante, vamos a considerar una clase de variables aleatorias algo más restrictivas en las que se da la igualdad.

Definición 4.2 (Estrictamente Subgaussiana).

Una variable aleatoria X se dice que es Estrictamente Subgaussiana si $X \sim \text{Sub}(\sigma^2)$ y $\sigma^2 = \mathbb{E}[X^2]$. Se denota $X \sim \text{SSub}(\sigma^2)$.

Cuando $X \sim \text{SSub}(\sigma^2)$, podemos obtener más información sobre sus momentos.

Proposición 4.4. Si $X \sim \text{SSub}(\sigma^2)$, entonces $\mathbb{E}[X^3] = 0$ y $\mathbb{E}[X^4] \leq 3\sigma^4$.

Demostración. Sabemos que

$$\mathbb{E}[e^{tX}] \leq e^{\frac{c^2 t^2}{2}},$$

y, como hacíamos en la otra proposición, escribiremos los desarrollos de las series, ahora hasta grado 4.

$$1 + t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}[X^2] + \frac{t^3}{3!}\mathbb{E}[X^3] + \frac{t^4}{4!}\mathbb{E}[X^4] + \mathcal{O}(|t|^5) \leq 1 + \frac{t^2\sigma^2}{2} + \frac{t^4\sigma^4}{8} + \mathcal{O}(|t|^5).$$

Empezamos calculando $\mathbb{E}[X^3]$, por lo que la expresión nos queda

$$\frac{t^3}{3!}\mathbb{E}[X^3] \leq t^4 \left(\frac{\sigma^4}{8} - \frac{\mathbb{E}[X^4]}{4!} \right) + \mathcal{O}(|t|^5) = \mathcal{O}(|t|^4).$$

- Para $t > 0$, dividiendo por $\frac{t^3}{3!}$, $\mathbb{E}[X^3] \leq \mathcal{O}(|t|)$ y tomando $t \rightarrow 0^+$ tenemos $\mathbb{E}[X^3] \leq 0$.
- Para $t < 0$, dividiendo por $\frac{t^3}{3!}$, $\mathbb{E}[X^3] \geq \mathcal{O}(|t|)$ y tomando $t \rightarrow 0^-$ tenemos $\mathbb{E}[X^3] \geq 0$.

Por tanto, $\mathbb{E}[X^3] = 0$ y nos quedará $\frac{t^4}{4!}\mathbb{E}[X^4] \leq \frac{\sigma^4 t^4}{8} + \mathcal{O}(|t|^5)$.

Para $t \neq 0$, si dividimos por $\frac{t^4}{4!}$ llegamos a la relación $\mathbb{E}[X^4] \leq \frac{4!}{8}\sigma^4 + \mathcal{O}(|t|) = 3\sigma^4 + \mathcal{O}(|t|)$. Y, tomando $t \rightarrow 0$, entonces

$$\mathbb{E}[X^4] \leq 3\sigma^4.$$

□

En definitiva, hemos probado lo siguiente:

| | $\mathbb{E}[X]$ | $\mathbb{E}[X^2]$ | $\mathbb{E}[X^3]$ | $\mathbb{E}[X^4]$ |
|--------------------|-----------------|-------------------|-------------------|-------------------|
| Sub(c^2) | = 0 | $\leq c^2$ | | |
| SSub(σ^2) | = 0 | = σ^2 | = 0 | $\leq 3\sigma^4$ |

Ahora ya podemos pasar a probar la propiedad de Concentración de Medida.

2. Concentración de Medida

Para esta sección, excepto algún lema, nos basaremos en el capítulo 7 del libro [3]. Antes de trabajar con las matrices aleatorias nos centraremos en los vectores aleatorios generados por una distribución Subgaussiana.

El teorema principal al que queremos llegar es el siguiente, pues la propiedad de Concentración de Medida será un corolario suyo.

Teorema 4.5. *Sea $X = [X_1, \dots, X_m]$, donde los $\{X_i\}_{i=1}^m$ son i.i.d. con $X_i \sim \text{SSub}(\sigma^2)$, entonces $\mathbb{E}[\|X\|_2^2] = m\sigma^2$. Además, si fijamos $\beta_{\text{máx}} > 1$, existirá una constante $K^* \geq 6$ dependiendo solo de β_{max} tal que se cumple*

$$\mathbb{P}(\|X\|_2^2 \leq \alpha m\sigma^2) \leq e^{-m(1-\alpha)^2/K^*}, \text{ para cualquier } \alpha \in (0, 1), \quad (4.1)$$

y

$$\mathbb{P}(\|X\|_2^2 \geq \beta m\sigma^2) \leq e^{-m(\beta-1)^2/K^*}, \text{ para cualquier } \beta \in [1, \beta_{\text{máx}}]. \quad (4.2)$$

Lo probaremos en dos partes, primero la desigualdad (4.1) del α y luego la (4.2) del β . En ambos casos usaremos muy a menudo la llamada Desigualdad de Markov:

Lema 4.6 (Desigualdad de Markov).

Para cualquier variable aleatoria no negativa X y $t > 0$, $\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$.

Demostración. Llamamos $A := \{\omega \in \Omega : X(\omega) \geq t\}$. Entonces,

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) d\mathbb{P}(\omega) \geq \int_A X(\omega) d\mathbb{P}(\omega) \geq t \int_A d\mathbb{P}(\omega) = t\mathbb{P}(A) = t\mathbb{P}(X \geq t).$$

□

Para probar la desigualdad (4.1) del α nos hará falta el siguiente lema, que podemos encontrar en [14]:

Lema 4.7. *Si $X \sim \text{SSub}(\sigma^2)$, entonces $\mathbb{E}[e^{-\lambda X^2}] \leq e^{-\lambda\sigma^2 + \frac{3}{2}\lambda^2\sigma^4}$, $\forall \lambda > 0$.*

Demostración. Usaremos estas propiedades de la exponencial:

$$a) \ e^{-x} \leq 1 - x + \frac{x^2}{2}, \ \forall x \geq 0. \quad b) \ 1 - x \leq e^{-x}, \ \forall x \in \mathbb{R}.$$

- Consideramos $b(x) := e^{-x} - (1 - x)$.

Tenemos $b'(x) = -e^{-x} + 1 = 0 \iff x = 0$ y como $b''(x) = e^{-x} > 0$, entonces $x = 0$ es el mínimo de $b(x)$. Así, como $b(0) = 0$ tendremos $b(x) \geq 0$, $\forall x \in \mathbb{R}$ y, por tanto, $1 - x \leq e^{-x}$, $\forall x \in \mathbb{R}$.

- Consideramos $a(x) := 1 - x + \frac{x^2}{2} - e^{-x}$.

Tenemos $a'(x) = -1 + x + e^{-x} = e^{-x} - (1 - x) = b(x) \geq 0$, $\forall x \in \mathbb{R}$. Luego, como $a(x)$ es creciente y $a(0) = 0$ tendremos $a(x) \geq 0$, $\forall x \geq 0$, y entonces $e^{-x} \leq 1 - x + \frac{x^2}{2}$, $\forall x \geq 0$.

Así,

$$\begin{aligned} \mathbb{E}[e^{-\lambda X^2}] &\stackrel{(a)}{\leq} \mathbb{E}\left[1 - \lambda X^2 + \frac{(\lambda X^2)^2}{2}\right] = 1 - \lambda \mathbb{E}[X^2] + \frac{\lambda^2}{2} \mathbb{E}[X^4] \leq \leftarrow \begin{cases} \mathbb{E}[X^2] = \sigma^2 \\ \mathbb{E}[X^4] \leq 3\sigma^4 \end{cases} \\ &\leq 1 - \lambda\sigma^2 + \frac{3}{2}\lambda^2\sigma^4 \stackrel{(b)}{\leq} e^{-\lambda\sigma^2 + \frac{3}{2}\lambda^2\sigma^4}. \end{aligned}$$

□

Teorema 4.8. Si $X = [X_1, \dots, X_m]$ donde los $\{X_i\}_{i=1}^m$ son i.i.d. con $X_i \sim \text{SSub}(\sigma^2)$, entonces $\forall \alpha \in (0, 1)$ se tiene

$$\mathbb{P}(\|X\|_2^2 \leq \alpha m \sigma^2) \leq e^{-m(1-\alpha)^2/6}.$$

Demostración. Para $\lambda > 0$ hacemos

$$\begin{aligned} \mathbb{P}(\|X\|_2^2 \leq \alpha m \sigma^2) &= \mathbb{P}(e^{-\lambda \|X\|_2^2} \geq e^{-\lambda \alpha m \sigma^2}) \stackrel{(\text{Markov})}{\leq} e^{\lambda \alpha m \sigma^2} \mathbb{E}[e^{-\lambda \|X\|_2^2}] \stackrel{(\text{ind})}{=} \\ &= e^{\lambda \alpha m \sigma^2} \prod_{i=1}^m \mathbb{E}[e^{-\lambda X_i^2}] \stackrel{(\text{Lema 4.7})}{\leq} e^{\lambda \alpha m \sigma^2} \prod_{i=1}^m e^{-\lambda \sigma^2 + \frac{3}{2}\lambda^2 \sigma^4} = \\ &= \left(e^{\lambda \alpha \sigma^2} e^{-\lambda \sigma^2 + \frac{3}{2}\lambda^2 \sigma^4}\right)^m = \left(e^{-(1-\alpha)\lambda \sigma^2 + \frac{3}{2}\lambda^2 \sigma^4}\right)^m. \end{aligned}$$

Si escogemos $\lambda = (1 - \alpha)\beta > 0$, con $\beta > 0$, tenemos

$$-(1 - \alpha)\lambda \sigma^2 + \frac{3}{2}\lambda^2 \sigma^4 = -(1 - \alpha)^2 \beta \sigma^2 + \frac{3}{2}(1 - \alpha)^2 \beta^2 \sigma^4 = -(1 - \alpha)^2 \left[\beta \sigma^2 - \frac{3}{2}\beta^2 \sigma^4\right].$$

Nos queda entonces

$$\mathbb{P}(\|X\|_2^2 \leq \alpha m \sigma^2) \leq \left(e^{-(1-\alpha)^2 \left[\beta \sigma^2 - \frac{3}{2}\beta^2 \sigma^4\right]}\right)^m$$

y si calculamos el ínfimo

$$\inf_{\beta > 0} \left(e^{-(1-\alpha)^2 \left[\beta \sigma^2 - \frac{3}{2}\beta^2 \sigma^4\right]}\right)^m = \left(e^{-(1-\alpha)^2 \sup_{\beta > 0} \left[\beta \sigma^2 - \frac{3}{2}\beta^2 \sigma^4\right]}\right)^m.$$

Vamos a calcular el óptimo. Para ello definimos $f(\beta) = \beta \sigma^2 - \frac{3}{2}\beta^2 \sigma^4$.

$$f'(\beta) = \sigma^2 - 3\beta \sigma^4 = 0 \iff \beta = \frac{\sigma^2}{3\sigma^4} = \frac{1}{3\sigma^2} > 0$$

que es el máximo pues $f''(\beta) = -3\sigma^4 < 0$. Así, si sustituimos este valor tendremos

$$\beta \sigma^2 - \frac{3}{2}\beta^2 \sigma^4 = \frac{\sigma^2}{3\sigma^2} - \frac{3}{2} \frac{\sigma^4}{9\sigma^4} = \frac{1}{3} - \frac{1}{6} = \frac{1}{6}.$$

Por tanto,

$$\mathbb{P}(\|X\|_2^2 \leq \alpha m \sigma^2) \leq e^{-m(1-\alpha)^2/6} \text{ para } \alpha \in (0, 1).$$

□

Con esto ya tendríamos probada la desigualdad (4.1) del α . Ahora vamos a por la (4.2) de β . Para ello, otra vez, nos hará falta la ayuda de un par de lemas.

Lema 4.9. *Sea $a \in \mathbb{R}$ y $b > 0$, entonces*

$$\int_{-\infty}^{+\infty} e^{at-bt^2} dt = \sqrt{\frac{\pi}{b}} e^{\frac{a^2}{4b}}.$$

Demostración. Sabemos que si $Z \sim \text{Normal}(0,1)$, entonces su función de densidad es $f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ y cumplirá

$$\int_{-\infty}^{+\infty} f(z) dz = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1.$$

Empezaremos completando cuadrados:

$$\left(\sqrt{bt} - \frac{a}{2\sqrt{b}} \right)^2 = bt^2 - at + \frac{a^2}{4b} \implies - \left(\sqrt{bt} - \frac{a}{2\sqrt{b}} \right)^2 + \frac{a^2}{4b} = -bt^2 + at$$

Luego,

$$\int_{-\infty}^{+\infty} e^{at-bt^2} dt = \int_{-\infty}^{+\infty} e^{-\left(\sqrt{bt} - \frac{a}{2\sqrt{b}}\right)^2 + \frac{a^2}{4b}} dt = e^{\frac{a^2}{4b}} \int_{-\infty}^{+\infty} e^{-\left(\sqrt{bt} - \frac{a}{2\sqrt{b}}\right)^2} dt.$$

Hacemos el cambio de variable $\frac{1}{\sqrt{2}}z = \sqrt{bt} + \frac{a}{2\sqrt{b}}$ con lo que $\frac{1}{\sqrt{2b}}dz = dt$ y

$$\int_{-\infty}^{+\infty} e^{at-bt^2} dt = e^{\frac{a^2}{4b}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2b}} e^{-\frac{z^2}{2}} dz = \sqrt{\frac{\pi}{b}} e^{\frac{a^2}{4b}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \sqrt{\frac{\pi}{b}} e^{\frac{a^2}{4b}}.$$

□

Lema 4.10. *Si $X \sim \text{Sub}(c^2)$, entonces $\mathbb{E}\left[e^{\frac{\lambda X^2}{2c^2}}\right] \leq \frac{1}{\sqrt{1-\lambda}}$, $\forall \lambda \in [0,1)$.*

Demostración. El caso $\lambda = 0$ es trivial, así que, suponemos que $\lambda \in (0,1)$.

Como $X \sim \text{Sub}(c^2)$, por definición

$$\mathbb{E}\left[e^{tX}\right] = \int_{\Omega} e^{tX(\omega)} d\mathbb{P}(\omega) \leq e^{\frac{c^2 t^2}{2}}, \text{ para cualquier } t \in \mathbb{R}.$$

Multiplicando por $e^{-\frac{c^2 t^2}{2\lambda}}$ e integrando respecto de t , obtenemos

$$\int_{-\infty}^{+\infty} \left(\int_{\Omega} e^{tX(\omega) - \frac{c^2 t^2}{2\lambda}} d\mathbb{P}(\omega) \right) dt \leq \int_{-\infty}^{+\infty} e^{\frac{c^2 t^2}{2\lambda}(\lambda-1)} dt.$$

Por un lado, cambiando el orden de integración tenemos

$$\begin{aligned} \int_{-\infty}^{+\infty} \left(\int_{\Omega} e^{tX(\omega) - \frac{c^2 t^2}{2\lambda}} d\mathbb{P}(\omega) \right) dt &= \int_{\Omega} \left(\int_{-\infty}^{+\infty} e^{tX(\omega) - \frac{c^2 t^2}{2\lambda}} dt \right) d\mathbb{P}(\omega) \stackrel{\text{(Lema 4.9)}}{=} \\ &= \int_{\Omega} \frac{\sqrt{2\pi\lambda}}{c} e^{\frac{\lambda X(\omega)^2}{2c^2}} d\mathbb{P}(\omega) = \frac{\sqrt{2\pi\lambda}}{c} \int_{\Omega} e^{\frac{\lambda X(\omega)^2}{2c^2}} d\mathbb{P}(\omega) = \frac{\sqrt{2\pi\lambda}}{c} \mathbb{E}\left[e^{\frac{\lambda X^2}{2c^2}}\right]. \end{aligned}$$

Por otro lado, volviendo a usar el Lema 4.9, se llega a

$$\int_{-\infty}^{+\infty} e^{\frac{c^2 t^2}{2\lambda}(\lambda-1)} dt = \sqrt{\frac{2\pi\lambda}{c^2(1-\lambda)}} = \frac{\sqrt{2\pi\lambda}}{c} \frac{1}{\sqrt{1-\lambda}}.$$

Por tanto,

$$\mathbb{E}\left[e^{\frac{\lambda X^2}{2c^2}}\right] \leq \frac{1}{\sqrt{1-\lambda}}, \quad \text{para } \lambda \in [0, 1].$$

□

Teorema 4.11. Si $X = [X_1, \dots, X_m]$ donde los $\{X_i\}_{i=1}^m$ son i.i.d. con $X_i \sim \text{SSub}(\sigma^2)$ y fijamos un $\beta_{\text{máx}} > 1$, entonces existe una constante K dependiendo solo de $\beta_{\text{máx}}$ tal que para cualquier $\beta \in [1, \beta_{\text{máx}}]$ se cumple

$$\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) \leq e^{-m(\beta-1)^2/K}.$$

Demostración. Para $\beta \geq 1$ y $\lambda > 0$ tenemos

$$\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) = \mathbb{P}(e^{\lambda \|X\|_2^2} \geq e^{\lambda \beta m \sigma^2}) \stackrel{\text{(Markov)}}{\leq} \frac{\mathbb{E}[e^{\lambda \|X\|_2^2}]}{e^{\lambda \beta m \sigma^2}} \stackrel{\text{(ind)}}{=} \frac{\prod_{i=1}^m \mathbb{E}[e^{\lambda X_i^2}]}{e^{\lambda \beta m \sigma^2}}.$$

Como $X_i \sim \text{Sub}(\sigma^2)$, por el Lema 4.10,

$$\mathbb{E}[e^{\lambda X_i^2}] = \mathbb{E}\left[e^{2\sigma^2 \lambda \frac{X_i^2}{2\sigma^2}}\right] \leq \frac{1}{\sqrt{1-2\sigma^2\lambda}} \quad \text{para } 2\sigma^2\lambda \in [0, 1) \iff \lambda \in [0, \frac{1}{2\sigma^2}).$$

Luego, $\prod_{i=1}^m \mathbb{E}[e^{\lambda X_i^2}] \leq \left(\frac{1}{\sqrt{1-2\sigma^2\lambda}}\right)^m = \left(\frac{1}{1-2\sigma^2\lambda}\right)^{m/2}$ y entonces

$$\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) \leq \frac{1}{e^{\lambda \beta m \sigma^2}} \left(\frac{1}{1-2\sigma^2\lambda}\right)^{m/2} = \left(\frac{e^{-2\lambda\beta\sigma^2}}{1-2\sigma^2\lambda}\right)^{m/2}, \quad \lambda \in [0, \frac{1}{2\sigma^2}).$$

Ahora, si minimizamos respecto de $\lambda \in [0, \frac{1}{2\sigma^2})$, se cumple

$$\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) \leq \inf_{\lambda \in [0, \frac{1}{2\sigma^2})} \left(\frac{e^{-2\lambda\beta\sigma^2}}{1-2\sigma^2\lambda}\right)^{m/2} = \left(\inf_{\lambda \in [0, \frac{1}{2\sigma^2})} \frac{e^{-2\lambda\beta\sigma^2}}{1-2\sigma^2\lambda}\right)^{m/2}.$$

Como tenemos $\frac{e^{-2\lambda\beta\sigma^2}}{1-2\sigma^2\lambda}$, $\lambda \in [0, \frac{1}{2\sigma^2})$, consideramos $g(t) := \frac{e^{-\beta t}}{1-t}$, $t \in [0, 1)$.

Vamos a calcular el óptimo.

$$\begin{aligned} g'(t) &= \frac{-\beta e^{-\beta t}(1-t) + e^{-\beta t}}{(1-t)^2} = \frac{e^{-\beta t}}{(1-t)^2} [-\beta(1-t) + 1] = -\frac{\beta e^{-\beta t}}{(1-t)^2} \left[(1-t) - \frac{1}{\beta}\right] = \\ &= -\frac{\beta e^{-\beta t}}{(1-t)^2} \left[\left(1 - \frac{1}{\beta}\right) - t\right] = 0 \iff t = 1 - \frac{1}{\beta} \in [0, 1) \text{ ya que } \beta \geq 1. \end{aligned}$$

Entonces el ínfimo se alcanza en $t = 1 - \frac{1}{\beta} = 2\sigma^2\lambda$, es decir, en $\lambda = \frac{1-\frac{1}{\beta}}{2\sigma^2} \in [0, \frac{1}{2\sigma^2})$.

Sustituyendo ese valor se obtiene

$$\frac{e^{-2\lambda\beta\sigma^2}}{1-2\sigma^2\lambda} = \frac{e^{-\beta(1-\frac{1}{\beta})}}{1-(1-\frac{1}{\beta})} = \beta e^{-(\beta-1)}.$$

Nos queda entonces, $\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) \leq (\beta e^{-(\beta-1)})^{m/2}$, con $\beta \geq 1$.

Para $\beta = 1$, el teorema se cumple de manera trivial. Por ello, nos centraremos en $\beta > 1$. Se puede ver que la función $h(\beta) = \frac{2(\beta-1)^2}{(\beta-1) - \log(\beta)}$ es creciente para $\beta > 1$. Así, dado $\beta_{\max} > 1$ si tomamos $K := h(\beta_{\max})$ tendremos que, para $\beta \in (1, \beta_{\max}]$,

$$\frac{2(\beta-1)^2}{(\beta-1) - \log(\beta)} \leq K \quad \text{luego} \quad \log(\beta) \leq (\beta-1) - \frac{2(\beta-1)^2}{K}$$

$$\text{y, por último,} \quad \beta \leq e^{(\beta-1) - \frac{2(\beta-1)^2}{K}}.$$

Por tanto, para $\beta \in (1, \beta_{\max}]$ se tiene

$$\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) \leq (\beta e^{-(\beta-1)})^{m/2} \leq \left(e^{(\beta-1) - \frac{2(\beta-1)^2}{K}} e^{-(\beta-1)} \right)^{m/2} = e^{-m(\beta-1)^2/K}.$$

□

Observación. Para el Lema 4.7, que usábamos para probar la desigualdad del α , si que era necesario que la variable fuera Estrictamente Subgaussiana pero para el Lema 4.10, que es el se usa para la desigualdad del β , nos basta con que sea Subgaussiana.

Y es que en la prueba original de [3] la desigualdad del β se hace directamente para Subgaussianas. La diferencia es que entonces tendríamos $X \sim \text{Sub}(c^2)$ y $\sigma^2 = \mathbb{E}[X^2] \leq c^2$, con lo que $1 \leq \frac{c^2}{\sigma^2} \leq \beta$ y habría que hacer la prueba con más cuidado.

Ahora ya podemos combinar los Teoremas 4.8 del α y 4.11 del β para probar el resultado que queríamos.

Teorema 4.12. Sea $X = [X_1, \dots, X_m]$, donde los $\{X_i\}_{i=1}^m$ son i.i.d. con $X_i \sim \text{SSub}(\sigma^2)$, entonces $\mathbb{E}[\|X\|_2^2] = m\sigma^2$. Además, si fijamos $\beta_{\max} > 1$, existirá una constante $K^* \geq 6$ dependiendo solo de β_{\max} tal que se cumple

$$\mathbb{P}(\|X\|_2^2 \leq \alpha m \sigma^2) \leq e^{-m(1-\alpha)^2/K^*}, \quad \text{para cualquier } \alpha \in (0, 1), \quad (4.3)$$

y

$$\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) \leq e^{-m(\beta-1)^2/K^*}, \quad \text{para cualquier } \beta \in [1, \beta_{\max}]. \quad (4.4)$$

Demostración. Empezamos viendo que $\mathbb{E}[\|X\|_2^2] = m\sigma^2$,

$$\mathbb{E}[\|X\|_2^2] = \mathbb{E}\left[\sum_{i=1}^m X_i^2\right] = \sum_{i=1}^m \mathbb{E}[X_i^2] = \sum_{i=1}^m \sigma^2 = m\sigma^2.$$

Por un lado, como $\alpha \in (0, 1)$, sabemos por el Teorema 4.8 que

$$\mathbb{P}(\|X\|_2^2 \leq \alpha m \sigma^2) \leq e^{-m(1-\alpha)^2/6}.$$

Por otro, para $\beta \in [1, \beta_{\max}]$, por el Teorema 4.11 tenemos

$$\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) \leq e^{-m(\beta-1)^2/K}, \quad \text{donde } K = \frac{2(\beta_{\max} - 1)^2}{(\beta_{\max} - 1) - \log(\beta_{\max})}.$$

Entonces, si tomamos $K^* := \max(6, K)$ obtendremos

$$\mathbb{P}(\|X\|_2^2 \leq \alpha m \sigma^2) \leq e^{-m(1-\alpha)^2/K^*}$$

y

$$\mathbb{P}(\|X\|_2^2 \geq \beta m \sigma^2) \leq e^{-m(\beta-1)^2/K^*}.$$

□

Como consecuencia del teorema anterior podemos probar la Concentración de Medida para vectores aleatorios.

Corolario 4.13. *Sea $X = [X_1, \dots, X_m]$, donde los $\{X_i\}_{i=1}^m$ son i.i.d. con $X_i \sim \text{SSub}(\sigma^2)$, entonces $\mathbb{E}[\|X\|_2^2] = m\sigma^2$ y para cualquier $\varepsilon \in (0, 1)$,*

$$\mathbb{P}(|\|X\|_2^2 - m\sigma^2| \geq \varepsilon m \sigma^2) \leq 2e^{-m\varepsilon^2/K^*}, \text{ con } K^* = \frac{2}{1 - \log(2)} \approx 6,52. \quad (4.5)$$

Demostración. Aplicamos el teorema anterior con $\alpha = 1 - \varepsilon$ y $\beta = 1 + \varepsilon$ para obtener

$$\mathbb{P}(\|X\|_2^2 \leq (1 - \varepsilon)m\sigma^2) \leq e^{-m\varepsilon^2/K^*}$$

y

$$\mathbb{P}(\|X\|_2^2 \geq (1 + \varepsilon)m\sigma^2) \leq e^{-m\varepsilon^2/K^*}.$$

Luego,

$$\begin{aligned} \mathbb{P}(|\|X\|_2^2 - m\sigma^2| \geq \varepsilon m \sigma^2) &= \mathbb{P}(\|X\|_2^2 - m\sigma^2 \leq -\varepsilon m \sigma^2) + \mathbb{P}(\|X\|_2^2 - m\sigma^2 \geq \varepsilon m \sigma^2) = \\ &= \mathbb{P}(\|X\|_2^2 \leq (1 - \varepsilon)m\sigma^2) + \mathbb{P}(\|X\|_2^2 \geq (1 + \varepsilon)m\sigma^2) \leq 2e^{-m\varepsilon^2/K^*}. \end{aligned}$$

Como $\varepsilon \in (0, 1)$, $\beta = 1 + \varepsilon \leq 2$ y podemos tomar $\beta_{\max} = 2$ de manera que la constante K^* nos quedará

$$K^* = \max\left(6, \frac{2(2-1)^2}{(2-1) - \log(2)}\right) = \max\left(6, \frac{2}{1 - \log(2)}\right) = \frac{2}{1 - \log(2)} \approx 6,52.$$

□

Para poder probar la Concentración de Medida para matrices aleatorias a partir de este corolario usaremos la siguiente proposición.

Proposición 4.14. *Suponemos que tenemos $X = [X_1, \dots, X_N]$ donde los $\{X_i\}_{i=1}^m$ son i.i.d. con $X_i \sim \text{SSub}(\sigma^2)$, entonces para cualquier $v \in \mathbb{R}^N$, $\langle X, v \rangle \sim \text{Sub}(c^2\|v\|_2^2)$.*

En particular, si cada $X_i \sim \text{SSub}(\sigma^2)$, entonces $\langle X, v \rangle \sim \text{SSub}(\sigma^2\|v\|_2^2)$.

Demostración. Se tiene

$$\begin{aligned} \mathbb{E}[e^{t\langle X, v \rangle}] &= \mathbb{E}\left[e^{t\sum_{i=1}^N v_i X_i}\right] = \mathbb{E}\left[\prod_{i=1}^N e^{tv_i X_i}\right] \stackrel{(\text{ind})}{=} \prod_{i=1}^N \mathbb{E}[e^{tv_i X_i}] \stackrel{\text{Sub}(c^2)}{\leq} \\ &\leq \prod_{i=1}^N \exp\left(\frac{c^2(tv_i)^2}{2}\right) = \exp\left(\frac{t^2}{2}c^2 \sum_{i=1}^N v_i^2\right) = \exp\left(\frac{t^2}{2}c^2\|v\|_2^2\right). \end{aligned}$$

Por tanto, $\langle X, v \rangle \sim \text{Sub}(c^2\|v\|_2^2)$.

Ahora, si $X_i \sim \text{SSub}(\sigma^2)$, ya tenemos $\langle X, v \rangle \sim \text{Sub}(\sigma^2 \|v\|_2^2)$. Para que sea Estrictamente Subgaussiana nos falta ver que $\mathbb{E}[\langle X, v \rangle^2] = \sigma^2 \|v\|_2^2$.

Al ser $X_i \sim \text{SSub}(\sigma^2)$ se cumple $\mathbb{E}[X_i^2] = \sigma^2$ y al ser independientes, como $\mathbb{E}[X_i] = 0$,

$$\mathbb{E}[X_i X_j] = \mathbb{E}[X_i] \mathbb{E}[X_j] = 0, \text{ para } i \neq j.$$

Así,

$$\begin{aligned} \mathbb{E}[\langle X, v \rangle^2] &= \mathbb{E} \left[\left(\sum_{i=1}^N v_i X_i \right) \left(\sum_{j=1}^N v_j X_j \right) \right] = \sum_{i,j=1}^N v_i v_j \mathbb{E}[X_i X_j] = \\ &= \sum_{i=1}^N v_i^2 \mathbb{E}[X_i^2] = \sigma^2 \sum_{i=1}^N v_i^2 = \sigma^2 \|v\|_2^2. \end{aligned}$$

Y, por tanto, $\langle X, v \rangle \sim \text{SSub}(\sigma^2 \|v\|_2^2)$. □

Por fin, podremos probar que una matriz aleatoria generada a partir de una distribución Estrictamente Subgaussiana cumplirá la propiedad de Concentración de Medida.

Teorema 4.15. *Sea Φ una matriz $m \times N$ cuyas entradas $\Phi_{i,j}$ son independientes e idénticamente distribuidas con $\Phi_{i,j} \sim \text{SSub}(1/m)$. Entonces, para cualquier $\varepsilon \in (0, 1)$ y cualquier $x \in \mathbb{R}^N$, tendremos $\mathbb{E}[\|\Phi x\|_2^2] = \|x\|_2^2$ y*

$$\mathbb{P} \left(\left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| \geq \varepsilon \|x\|_2^2 \right) \leq 2e^{-m\varepsilon^2/K^*} \quad (4.6)$$

con $K^* = \frac{2}{1-\log(2)} \approx 6,52$.

Demostración. Denotamos como Φ_i la i -ésima fila de Φ . Como $\Phi_i = [\Phi_{i,1}, \dots, \Phi_{i,N}]$ con $\{\Phi_{i,j}\}_{j=1}^N$ i.i.d. con $\Phi_{i,j} \sim \text{SSub}(1/m)$, por la Proposición 4.14, sabemos

$$\langle \Phi_i, x \rangle \sim \text{SSub}(\|x\|_2^2/m).$$

Si llamamos $y = \Phi x \in \mathbb{R}^m$, entonces se tiene $y_i = \langle \Phi_i, x \rangle \sim \text{SSub}(\|x\|_2^2/m)$ y, además, los y_i serán independientes. De esta manera, podemos aplicar el Corolario 4.13 al vector $y = [y_1, \dots, y_m]$ con el $\varepsilon \in (0, 1)$ dado y obtenemos

$$\mathbb{E}[\|\Phi x\|_2^2] = \mathbb{E}[\|y\|_2^2] = m \frac{\|x\|_2^2}{m} = \|x\|_2^2$$

y

$$\mathbb{P} \left(\left| \|\Phi x\|_2^2 - \|x\|_2^2 \right| \geq \varepsilon \|x\|_2^2 \right) = \mathbb{P} \left(\left| \|y\|_2^2 - \|x\|_2^2 \right| \geq \varepsilon \|x\|_2^2 \right) \leq 2e^{-m\varepsilon^2/K^*}$$

con $K^* = \frac{2}{1-\log(2)} \approx 6,52$. □

3. Caracterización y recíproco

En esta sección vamos a dar una caracterización para las variables Subgaussianas y veremos un resultado que podría verse como el recíproco a la cota superior de la Concentración de Medida. La caracterización de las distribuciones la podemos encontrar en [15].

Teorema 4.16 (Caracterización de Subgaussianas).

Para una variable aleatoria X centrada, o sea, con $E[X] = 0$, son equivalentes:

- 1) Existe $c > 0$ tal que $X \sim \text{Sub}(c^2)$, es decir, tal que $\mathbb{E}[e^{tX}] \leq e^{\frac{c^2 t^2}{2}}$, $\forall t \in \mathbb{R}$.
- 2) Existen $b_1, b_2 > 0$ tales que $\mathbb{P}(|X| \geq \lambda) \leq b_1 e^{-b_2 \lambda^2}$, $\forall \lambda > 0$.
- 3) Existe $a > 0$ tal que $\mathbb{E}[e^{aX^2}] \leq 2$.

Demostración. Haremos la demostración de manera cíclica.

- $\boxed{(1) \implies (2)}$ Para $\lambda, t > 0$ se tiene

$$\mathbb{P}(X \geq \lambda) = \mathbb{P}(e^{tX} \geq e^{t\lambda}) \stackrel{(\text{Markov})}{\leq} \frac{\mathbb{E}[e^{tX}]}{e^{t\lambda}} \stackrel{\text{Sub}(c^2)}{\leq} e^{-t\lambda + \frac{c^2 t^2}{2}}.$$

Como la exponencial es estrictamente creciente, minimizar $e^{-t\lambda + \frac{c^2 t^2}{2}}$ es lo mismo que minimizar el exponente, $f(t) := -t\lambda + \frac{c^2 t^2}{2}$. Haciendo los cálculos, como $f'(t) = -\lambda + c^2 t$, el mínimo se alcanza en $t = \lambda/c^2 > 0$. Entonces,

$$-t\lambda + \frac{c^2 t^2}{2} = -\frac{\lambda^2}{c^2} + \frac{c^2 \lambda^2}{2c^4} = -\frac{\lambda^2}{c^2} + \frac{\lambda^2}{2c^2} = -\frac{\lambda^2}{2c^2}$$

y, por tanto, $\mathbb{P}(X \geq \lambda) \leq e^{-\frac{\lambda^2}{2c^2}}$. Análogamente, se puede ver que $\mathbb{P}(X \leq -\lambda) \leq e^{-\frac{\lambda^2}{2c^2}}$.

Así, $\forall \lambda > 0$,

$$\mathbb{P}(|X| \geq \lambda) = \mathbb{P}(X \geq \lambda) + \mathbb{P}(X \leq -\lambda) \leq 2e^{-\frac{\lambda^2}{2c^2}}.$$

Luego, se cumple (2) con $b_1 = 2 > 0$ y $b_2 = \frac{1}{2c^2} > 0$.

- $\boxed{(2) \implies (3)}$ Dado $a > 0$ podemos considerar la función $H(t) = e^{at^2} - 1$ para $t \geq 0$.

Sabemos que H es derivable, creciente y, además, $H(0) = 0$. Podemos entonces aplicar el Lema C.1, que se encuentra en el Apéndice C, y tendremos

$$\mathbb{E}[e^{aX^2}] = 1 + \int_0^{+\infty} 2ate^{at^2} \mathbb{P}(|X| > t) dt.$$

Por hipótesis, existen $b_1, b_2 > 0$ tales que $\mathbb{P}(|X| \geq \lambda) \leq b_1 e^{-b_2 \lambda^2}$, $\forall \lambda > 0$. Así, si ahora consideramos $0 < a < b_2$ podremos acotar la integral,

$$\begin{aligned} \int_0^{+\infty} 2ate^{at^2} \mathbb{P}(|X| > t) dt &\leq \int_0^{+\infty} 2ate^{at^2} \cdot b_1 e^{-b_2 \lambda^2} dt = \int_0^{+\infty} 2ab_1 t e^{-(b_2 - a)t^2} dt = \\ &= -\frac{ab_1}{b_2 - a} \int_0^{+\infty} -2(b_2 - a)t e^{-(b_2 - a)t^2} dt = -\frac{ab_1}{b_2 - a} \left[e^{-(b_2 - a)t^2} \right]_0^{+\infty} = \frac{ab_1}{b_2 - a}. \end{aligned}$$

Por tanto, nos queda $\mathbb{E}[e^{aX^2}] \leq 1 + \frac{ab_1}{b_2 - a}$. En particular, tomando $0 < a := \frac{b_2}{1 + b_1} < b_2$ tendremos $\frac{ab_1}{b_2 - a} = 1$ y entonces se cumple (3) pues $\mathbb{E}[e^{aX^2}] \leq 2$.

- $\boxed{(3) \implies (1)}$ Para cualquier $t \in \mathbb{R}$,

$$\mathbb{E}[e^{tX}] = \mathbb{E}\left[\sum_{n \geq 0} \frac{(tX)^n}{n!}\right] = \mathbb{E}[1] + \mathbb{E}[X] + \mathbb{E}\left[\sum_{n \geq 2} \frac{(tX)^n}{n!}\right].$$

Para $k \geq 2$, $\frac{(tX)^k}{k!} \leq \frac{|tX|^k}{k!} \leq \frac{t^2 X^2 |tX|^{k-2}}{2(k-2)!}$, y podemos escribir

$$\sum_{n \geq 2} \frac{(tX)^n}{n!} \leq \frac{t^2 X^2}{2} \sum_{n \geq 2} \frac{|tX|^{n-2}}{(n-2)!} = \frac{t^2 X^2}{2} \sum_{k \geq 0} \frac{|tX|^k}{k!} = \frac{t^2 X^2}{2} e^{|tX|}.$$

Así, $\mathbb{E}[e^{tX}] \leq 1 + \frac{t^2}{2} \mathbb{E}[X^2 e^{|tX|}]$. Ahora vamos a ver que $|tX| \leq \left(\frac{t^2}{2a} + \frac{aX^2}{2}\right)$.

Sabemos que $2AB \leq A^2 + B^2$ y, como $a > 0$, podemos tomar $A = \frac{|t|}{\sqrt{2a}}$ y $B = \frac{\sqrt{a}|X|}{\sqrt{2}}$ y tendremos

$$|tX| = 2 \frac{|t|}{\sqrt{2a}} \cdot \frac{\sqrt{a}|X|}{\sqrt{2}} \leq \left(\frac{|t|}{\sqrt{2a}}\right)^2 + \left(\frac{\sqrt{a}|X|}{\sqrt{2}}\right)^2 = \frac{t^2}{2a} + \frac{aX^2}{2}.$$

Podemos entonces seguir con el desarrollo y llegar a

$$\mathbb{E}[e^{tX}] \leq 1 + \frac{t^2}{2} \mathbb{E}[X^2 e^{|tX|}] \leq 1 + \frac{t^2}{2} \mathbb{E}[X^2 e^{\frac{t^2}{2a} + \frac{aX^2}{2}}] = 1 + \frac{t^2}{2} e^{\frac{t^2}{2a}} \mathbb{E}[X^2 e^{\frac{aX^2}{2}}].$$

Para continuar usaremos la hipótesis $\mathbb{E}[e^{aX^2}] \leq 2$ y las propiedades de la exponencial:

$$(i) : z \leq e^{z/2} \quad \text{y} \quad (ii) : 1 + z \leq e^z.$$

Entonces

$$\begin{aligned} \mathbb{E}[e^{tX}] &\leq 1 + \frac{t^2}{2a} e^{\frac{t^2}{2a}} \mathbb{E}[aX^2 e^{\frac{aX^2}{2}}] \stackrel{(i)}{\leq} 1 + \frac{t^2}{2a} e^{\frac{t^2}{2a}} \mathbb{E}[e^{aX^2}] \stackrel{\text{hip.}}{\leq} 1 + \frac{t^2}{a} e^{\frac{t^2}{2a}} \leq \\ &\leq \left(1 + \frac{t^2}{a}\right) e^{\frac{t^2}{2a}} \stackrel{(ii)}{\leq} e^{\frac{t^2}{a}} e^{\frac{t^2}{2a}} = e^{\frac{3t^2}{2a}} = e^{\frac{t^2}{2} \cdot \frac{3}{a}}. \end{aligned}$$

Por tanto, $X \sim \text{Sub}(\frac{3}{a})$ y se cumple (1). □

Usando la caracterización, en concreto, el apartado (2) podremos probar el siguiente teorema, basado en el Lema 5.3 del artículo [10].

Teorema 4.17. *Sea X una variable aleatoria con $\mathbb{E}[X] = 0$ y $\mathbb{E}[X^2] = \sigma^2$. Si existe una constante $a > 0$ tal que para todo $m \in \mathbb{N}$ y $\vec{X} = [X_1, \dots, X_m]$, con $\{X_i\}_{i=1}^m$ i.i.d. como X , se cumple*

$$\mathbb{P}\left(\|\vec{X}\|_2^2 - \mathbb{E}[\|\vec{X}\|_2^2] \geq \frac{1}{2}m\sigma^2\right) \leq 2e^{-ma},$$

entonces $X \sim \text{Sub}(c^2)$, para algún $c > 0$.

Demostración. Por la caracterización de las Subgaussianas, Teorema 4.16, bastará probar que existen $b_1, b_2 > 0$ tales que

$$\mathbb{P}(|X| \geq \lambda) \leq b_1 e^{-b_2 \lambda^2}, \quad \forall \lambda > 0.$$

Dado $\lambda \geq \sqrt{\frac{3}{2}}\sigma > 0$, tomamos $m \in \mathbb{N}$ como $m := \left\lfloor \frac{\lambda^2}{\frac{3}{2}\sigma^2} \right\rfloor$ de manera que se cumple $\frac{3}{2}m\sigma^2 \leq \lambda^2 < \frac{3}{2}(m+1)\sigma^2$. Si consideramos $\vec{X} = [X_1, \dots, X_m]$ con $\{X_i\}_{i=1}^m$ i.i.d. como X tendremos, como ya hemos probado otras veces, que $\mathbb{E}[\|\vec{X}\|_2^2] = m\sigma^2$ con lo cual

$$\begin{aligned} \mathbb{P}(|X| \geq \lambda) &= \mathbb{P}(|X_1| \geq \lambda) = \mathbb{P}(|X_1|^2 \geq \lambda^2) \leq \mathbb{P}(\|\vec{X}\|_2^2 \geq \lambda^2) \leq \\ &\leq \mathbb{P}\left(\|\vec{X}\|_2^2 \geq \frac{3}{2}m\sigma^2\right) = \mathbb{P}\left(\|\vec{X}\|_2^2 - \mathbb{E}[\|\vec{X}\|_2^2] \geq \frac{1}{2}m\sigma^2\right). \end{aligned}$$

Y, usando la hipótesis, como $\lambda^2 < \frac{3}{2}(m+1)\sigma^2$, entonces $m \geq \frac{2\lambda^2}{3\sigma^2} - 1$ y

$$\mathbb{P}(|X| \geq \lambda) \leq \mathbb{P}\left(\|\vec{X}\|_2^2 - \mathbb{E}[\|\vec{X}\|_2^2] \geq \frac{1}{2}m\sigma^2\right) \leq 2e^{-ma} \leq 2e^{-a\left(\frac{2\lambda^2}{3\sigma^2} - 1\right)} = 2e^a e^{-a\frac{2\lambda^2}{3\sigma^2}}.$$

Podemos llamar $b_1 := 2e^a > 0$ y $b_2 := \frac{2a}{3\sigma^2} > 0$, de forma que $\forall \lambda \geq \sqrt{\frac{3}{2}}\sigma$,

$$\mathbb{P}(|X| \geq \lambda) \leq b_1 e^{-b_2 \lambda^2}.$$

Para los $0 < \lambda < \sqrt{\frac{3}{2}}\sigma$, tendremos

$$\mathbb{P}(|X| \geq \lambda) \leq 1 = e^{\frac{2\lambda^2}{3\sigma^2}a} e^{-\frac{2\lambda^2}{3\sigma^2}a} \leq e^a e^{-\frac{2\lambda^2}{3\sigma^2}a} \leq b_1 e^{-b_2 \lambda^2}.$$

Por tanto, como se cumple $\forall \lambda > 0$, existirá $c > 0$ tal que $X \sim \text{Sub}(c^2)$. □

Cuando decimos que una variable aleatoria X con $\mathbb{E}[X] = 0$ y $\mathbb{E}[X^2] = \sigma^2$ cumple la cota superior de la Concentración de Medida queremos decir que si generamos un vector aleatorio \vec{X} con ella de tamaño cualquiera $m \in \mathbb{N}$ entonces para cualquier $\varepsilon \in (0, 1)$ existe una constante $c_0(\varepsilon) > 0$ tal que

$$\mathbb{P}\left(\|\vec{X}\|_2^2 - \mathbb{E}[\|\vec{X}\|_2^2] \geq \varepsilon m\sigma^2\right) \leq 2e^{-mc_0(\varepsilon)}.$$

Recordamos que en la sección anterior decíamos que el Teorema 4.11 de la desigualdad del β podía probarse directamente para Subgaussianas. De esta manera, aunque el Teorema de Concentración de Medida lo enunciamos para Estrictamente Subgaussianas realmente la cota superior la cumplirán todas las Subgaussianas.

Teniendo esto en cuenta, el Teorema 4.17, que acabamos de probar, puede verse como un recíproco de la cota superior de la Concentración de Medida. Ya que si se cumple la desigualdad para cualquier $\varepsilon \in (0, 1)$ se cumplirá para $\varepsilon = 1/2$ y por el teorema la variable se distribuirá como una Subgaussiana.

4. Lema de Johnson-Lindenstrauss

Antes de pasar a ver algunos ejemplos de distribuciones Subgaussianas vamos a ver que se puede probar el célebre Lema de Johnson-Lindenstrauss como corolario al Teorema de Concentración de Medida.

Para este lema existen numerosas versiones, cada una con un enunciado diferente ya que las constantes varían según el método de demostración utilizado. En nuestro caso, la idea para la demostración la podemos encontrar en [2] y seguirá un desarrollo muy similar al de la prueba de RIP, ver Teorema 3.6, ya que como podemos observar ambos enunciados son muy similares.

Teorema 4.18. *Dado $0 < \varepsilon < 1$. Para todo conjunto de puntos Q en \mathbb{R}^N , si m es un entero positivo tal que $m > 3K^* \log(|Q|)/\varepsilon^2$, donde $K^* = \frac{2}{1-\log(2)} \approx 6,52$, entonces existe una aplicación lineal $f : \mathbb{R}^N \rightarrow \mathbb{R}^m$ tal que*

$$(1 - \varepsilon)\|u - v\|_2^2 \leq \|f(u) - f(v)\|_2^2 \leq (1 + \varepsilon)\|u - v\|_2^2, \quad \forall u, v \in Q. \quad (4.7)$$

Demostración. Para probarlo usaremos el Teorema 4.15 de Concentración de Medida para Estrictamente Subgaussianas.

Consideramos una matriz aleatoria Φ de tamaño $m \times N$ donde las $\Phi_{i,j}$ son i.i.d como $\Phi_{i,j} \sim \text{SSub}(1/m)$. Dados $u, v \in \mathbb{R}^N$, denotamos

$$\Omega(u, v) = \left\{ \omega \in \Omega : \left| \|\Phi(\omega)(u - v)\|_2^2 - \|u - v\|_2^2 \right| \leq \varepsilon \|u - v\|_2^2 \right\}.$$

Por el teorema sabemos que

$$\mathbb{P}(\Omega(u, v)^c) = \mathbb{P}\left(\left| \|\Phi(\omega)(u - v)\|_2^2 - \|u - v\|_2^2 \right| > \varepsilon \|u - v\|_2^2\right) \leq 2e^{-m\varepsilon^2/K^*}.$$

De esta manera, tendremos

$$\begin{aligned} \mathbb{P}\left(\left| \|\Phi(\omega)(u - v)\|_2^2 - \|u - v\|_2^2 \right| \leq \varepsilon \|u - v\|_2^2, \quad \forall u, v \in Q\right) &= \mathbb{P}\left(\bigcap_{u, v \in Q} \Omega(u, v)\right) = \\ &= 1 - \mathbb{P}\left(\bigcup_{u, v \in Q} \Omega(u, v)^c\right) \geq 1 - \sum_{u, v \in Q} \mathbb{P}(\Omega(u, v)^c) \geq 1 - \sum_{u, v \in Q} 2e^{-m\varepsilon^2/K^*} = \\ &= 1 - 2|Q|^2 e^{-m\varepsilon^2/K^*} > 1 - 2|Q|^2 e^{-3 \log(|Q|)}, \quad \text{ya que } m > 3K^* \log(|Q|)/\varepsilon^2. \end{aligned}$$

- Si $|Q| = 1$ tendríamos $\mathbb{P}\left(\bigcap_{u, v \in Q} \Omega(u, v)\right) > -1$ pero al tener solamente un elemento el resultado se cumple trivialmente para cualquier $m \in \mathbb{N}^+$.
- Si $|Q| > 1$, es decir, si $|Q| \geq 2$ tendremos

$$\mathbb{P}\left(\bigcap_{u, v \in Q} \Omega(u, v)\right) > 1 - 2|Q|^2 e^{\log(|Q|^{-3})} = 1 - 2 \frac{|Q|^2}{|Q|^3} = 1 - \frac{2}{|Q|} \geq 1 - \frac{2}{2} = 0.$$

Por tanto, con probabilidad > 0 , $\Phi(\omega)$ cumple

$$(1 - \varepsilon)\|u - v\|_2^2 \leq \|\Phi(\omega)(u) - \Phi(\omega)(v)\|_2^2 \leq (1 + \varepsilon)\|u - v\|_2^2, \forall u, v \in Q.$$

Esto quiere decir que debe existir una aplicación lineal $f : \mathbb{R}^N \rightarrow \mathbb{R}^m$ tal que

$$(1 - \varepsilon)\|u - v\|_2^2 \leq \|f(u) - f(v)\|_2^2 \leq (1 + \varepsilon)\|u - v\|_2^2, \forall u, v \in Q.$$

□

5. Ejemplos

Como ya habíamos dicho, para acabar el capítulo veremos algunos ejemplos de distribuciones Subgaussianas. Realmente veremos variables aleatorias que son Estrictamente Subgaussianas y que, por tanto, podremos utilizar para generar las matrices de medición en la reconstrucción de señales.

1) Distribución Normal

Si $X \sim \text{Normal}(0, \sigma^2)$, recordamos que su función de densidad es

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}, \forall x \in \mathbb{R}.$$

Entonces podemos usar el Lema 4.9 y tendremos

$$\mathbb{E}[e^{tX}] = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} e^{tx} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} e^{tx - \frac{x^2}{2\sigma^2}} dx \stackrel{\text{Lema}}{=} e^{\frac{\sigma^2 t^2}{2}}.$$

Por tanto, $X \sim \text{Sub}(\sigma^2)$. Además, como $\mathbb{E}[X^2] = \sigma^2$, tendremos $X \sim \text{SSub}(\sigma^2)$.

2) Distribución Rademacher

Una variable aleatoria (discreta) X con esta distribución tiene, para $a > 0$, la siguiente función de probabilidad:

$$\mathbb{P}(X = x) = \begin{cases} 1/2 & \text{si } x = -a. \\ 1/2 & \text{si } x = a. \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces,

$$\mathbb{E}[e^{tX}] = e^{-ta}\mathbb{P}(X = -a) + e^{ta}\mathbb{P}(X = a) = \frac{e^{-ta} + e^{ta}}{2} = \cosh(ta).$$

Vamos a ver, utilizando sus desarrollos en serie, que $\cosh(z) \leq e^{\frac{z^2}{2}}$. Para ello, primero veamos que se cumple $2^n n! \leq (2n)!$, $\forall n \in \mathbb{N}$.

Para $n = 0$ y $n = 1$ está claro que se cumple. Ahora, para $n \geq 2$,

$$(2n)! = \underbrace{2n(2n-1)\dots(n+1)n!}_{n \text{ términos } > 2} \geq 2^n n!.$$

Así,

$$\cosh(z) = \frac{e^{-z} + e^z}{2} = \sum_{n \geq 0} \frac{z^{2n}}{(2n)!} \leq \sum_{n \geq 0} \frac{z^{2n}}{2^n n!} = \sum_{n \geq 0} \frac{\left(\frac{z^2}{2}\right)^n}{n!} = e^{\frac{z^2}{2}}.$$

Volviendo a la variable aleatoria X , tendremos $\mathbb{E}[e^{tX}] = \cosh(ta) \leq e^{\frac{t^2 a^2}{2}}$ y, por tanto, $X \sim \text{Sub}(a^2)$. Además,

$$\mathbb{E}[X^2] = (-a)^2 \mathbb{P}(X = -a) + (a)^2 \mathbb{P}(X = a) = \frac{a^2}{2} + \frac{a^2}{2} = a^2.$$

Con lo que $X \sim \text{SSub}(a^2)$.

3) Distribución Uniforme

Si $X \sim \text{Unif}(-a, a)$ con $a > 0$ su función de densidad será $f(x) = \frac{1}{2a}$ en $[-a, a]$ y 0 en el resto. Luego,

$$\mathbb{E}[e^{tX}] = \int_{-a}^a \frac{1}{2a} e^{tx} dx = \left[\frac{e^{tx}}{2at} \right]_{x=-a}^{x=a} = \frac{1}{2ta} (e^{ta} - e^{-ta}) = \frac{1}{ta} \sinh(ta).$$

Podemos usar el desarrollo en serie de \sinh y la desigualdad anterior $2^n n! \leq (2n)! \leq (2n+1)!$ para probar que $X \sim \text{Sub}(a^2)$,

$$\mathbb{E}[e^{tX}] = \frac{1}{ta} \sinh(ta) = \frac{1}{ta} \sum_{n \geq 0} \frac{(ta)^{2n+1}}{(2n+1)!} = \sum_{n \geq 0} \frac{(ta)^{2n}}{(2n+1)!} \leq \sum_{n \geq 0} \frac{(ta)^{2n}}{2^n n!} = e^{\frac{t^2 a^2}{2}}.$$

No obstante, si calculamos $\mathbb{E}[X^2]$ obtenemos

$$\mathbb{E}[X^2] = \int_{-a}^a \frac{1}{2a} x^2 dx = \left[\frac{x^3}{6a} \right]_{x=-a}^{x=a} = \frac{a^3}{6a} - \frac{(-a)^3}{6a} = \frac{a^2}{3} < a^2.$$

Por tanto, para ver que es Estrictamente Subgaussiana necesitaremos $X \sim \text{Sub}\left(\frac{a^2}{3}\right)$.

Vamos a probar de manera general que $\frac{\sinh(z)}{z} \leq e^{\frac{z^2}{6}}$, para todo $z \neq 0$.

Primero veamos usando inducción que $6^n n! \leq (2n+1)!$, $\forall n \in \mathbb{N}$. Los casos base $n = 0$ y $n = 1$ son evidentes. Y, si suponemos que es cierto para $n \geq 0$, es sencillo ver que se cumple para $n+1$, pues

$$\begin{aligned} (2(n+1)+1)! &= (2n+3)! = (2n+3)(2n+2)(2n+1)! \geq (2n+3)(2n+2)6^n n! = \\ &= \underbrace{2(2n+3)}_{\geq 6} 6^n (n+1)! \geq 6^{n+1} (n+1)!. \end{aligned}$$

Con esto tendremos lo que queríamos,

$$\frac{\sinh(z)}{z} = \sum_{n \geq 0} \frac{z^{2n}}{(2n+1)!} \leq \sum_{n \geq 0} \frac{z^{2n}}{6^n n!} = \sum_{n \geq 0} \left(\frac{z^2}{6}\right)^n \frac{1}{n!} = e^{\frac{z^2}{6}}, \quad \forall z \neq 0.$$

Volviendo a nuestra variable X nos quedará

$$\mathbb{E}[e^{tX}] = \frac{1}{ta} \sinh(ta) \leq e^{\frac{t^2 a^2}{6}} = e^{\frac{t^2}{2} \frac{a^2}{3}}.$$

Luego $X \sim \text{Sub}(\frac{a^2}{3})$ y, por tanto, $X \sim \text{SSub}(\frac{a^2}{3})$.

$$4) \text{ Distribución } \begin{cases} \mathbb{P}(X = \pm a) = \frac{1-s}{2} \\ \mathbb{P}(X = 0) = s \end{cases} \quad \text{donde } s \in [0, \frac{2}{3}].$$

Empezaremos viendo el valor de $\mathbb{E}[X^2]$,

$$\begin{aligned} \mathbb{E}[X^2] &= (-a)^2\mathbb{P}(X = -a) + 0^2\mathbb{P}(X = 0) + a^2\mathbb{P}(X = a) = \\ &= a^2\frac{1-s}{2} + a^2\frac{1-s}{2} = a^2(1-s). \end{aligned}$$

Entonces para ver que X es una variable Estrictamente Subgaussiana necesitaremos que $X \sim \text{Sub}(a^2(1-s))$.

Ahora vamos a calcular $\mathbb{E}[e^{tX}]$,

$$\begin{aligned} \mathbb{E}[e^{tX}] &= e^{-ta}\mathbb{P}(X = -a) + e^{t0}\mathbb{P}(X = 0) + e^{ta}\mathbb{P}(X = a) = \\ &= e^{-ta}\frac{1-s}{2} + s + e^{ta}\frac{1-s}{2} = \left(\frac{1-s}{2}\right)(e^{-ta} + e^{ta}) + s = \\ &= (1-s)\cosh(ta) + s \stackrel{?}{\leq} e^{\frac{a^2t^2}{2}(1-s)}. \end{aligned}$$

Consideramos $f(x) := e^{\frac{x^2}{2}(1-s)} - (1-s)\cosh(x) - s$. Queremos que $f(x) \geq 0, \forall x \in \mathbb{R}$. Para $x = 0, f(0) = 1 - (1-s) - s = 0$. Luego, si conseguimos ver que el 0 es el mínimo entonces tendremos la desigualdad probada. Calculemos su derivada,

$$f'(x) = e^{\frac{x^2}{2}(1-s)}x(1-s) - (1-s)\sinh(x) = (1-s)\left(e^{\frac{x^2}{2}(1-s)}x - \sinh(x)\right).$$

Para $x = 0, f'(0) = (1-s)(0-0) = 0$, y para $x \neq 0, f'(x) = (1-s)x\left(e^{\frac{x^2}{2}(1-s)} - \frac{\sinh(x)}{x}\right)$.

En el caso anterior de la Distribución Uniforme probamos que $\frac{\sinh(x)}{x} \leq e^{\frac{x^2}{6}}, \forall x \neq 0$. Con esa desigualdad y como estamos suponiendo que $s \in [0, \frac{2}{3}]$, tendremos

$$\begin{aligned} e^{\frac{x^2}{2}(1-s)} - \frac{\sinh(x)}{x} &\geq e^{\frac{x^2}{2}(1-s)} - e^{\frac{x^2}{6}} \geq 0 \iff \frac{x^2}{2}(1-s) \geq \frac{x^2}{6} \\ (1-s) &\geq \frac{1}{3} \iff \frac{2}{3} \geq s. \end{aligned}$$

Por tanto, el signo de la derivada solo depende del signo de x y 0 será el mínimo. Entonces se tiene lo que queríamos,

$$\mathbb{E}[e^{tX}] = (1-s)\cosh(ta) + s \leq e^{\frac{a^2t^2}{2}(1-s)}, \forall t \in \mathbb{R},$$

y $X \sim \text{SSub}(a^2(1-s))$.

Capítulo 5

Minimización ℓ_1 en Octave

En este capítulo trabajaremos con el software Octave [4] para poner a prueba el método de la Minimización ℓ_1 . Para disponer de este método en Octave, utilizaremos la función `l1eq_pd()` del paquete ℓ_1 -MAGIC [8]. Su implementación está basada en el algoritmo primal-dual de puntos interiores que se describe en el Capítulo 7 del libro [5] y, aunque no entraremos en más detalles sobre el algoritmo, sí que explicaremos sus distintos parámetros y cómo utilizarla.

Para llamar a la función haremos:

```
xp = l1eq_pd(x0, A, At, b, pdtol, pdmaxiter, cgtol, cgmaxiter)
```

donde:

- `x0` es el vector $N \times 1$ inicial con el que empezar el método.
- `A` es, o bien una función que toma vectores de tamaño N y devuelve vectores de tamaño m , o bien una matriz de tamaño $m \times N$.
- `At` es una función que toma vectores de tamaño m y devuelve vectores de tamaño N . Si `A` es una matriz de tamaño $m \times N$, entonces `At` es ignorada.
- `b` es el vector $m \times 1$ de observaciones.
- `pdtol` es la tolerancia para el algoritmo primal-dual, es decir, el algoritmo terminará si el salto de dualidad es menor que `pdtol`. Por defecto, `pdtol` es igual a 10^{-3} .
- `pdmaxiter` es el número máximo de iteraciones del algoritmo primal-dual. Por defecto, `pdmaxiter` es igual a 50.
- `cgtol` y `cgmaxiter` son la tolerancia y el número máximo de iteraciones pero para el método de Gradientes Conjugados que también incluye el algoritmo. Estos parámetros serán ignorados si `A` es una matriz de tamaño $m \times N$.

En nuestro caso `A` será la matriz de mediciones, o sea, Φ , por lo que los parámetros `At`, `cgtol` y `cgmaxiter` los podemos omitir. En cuanto a `x0` podremos usar cualquier vector inicial de tamaño correcto pues en caso de no ser un punto factible el algoritmo nos lo corregirá. En los ejemplos del propio paquete utilizan como vector inicial el llamado punto de mínima energía que no es más que $A' * b$. Realmente como los ejemplos que vamos a probar no son excesivamente complejos podremos omitir `pdtol` y `pdmaxiter` y dejar los valores por defecto. Y, por supuesto, para el vector de mediciones `b` usaremos $y = \Phi x$.

En definitiva, cuando llamemos a la función haremos simplemente:

```
xp = l1eq_pd(x0, Phi, [], y);
```

La matriz de mediciones la construiremos de manera aleatoria a partir de una distribución Estrictamente Subgaussiana para asegurarnos que se cumple la Concentración de Medida y que entonces con alta probabilidad la reconstrucción sea exacta. En concreto, usaremos una Normal(0, 1/ m). En el Apéndice D se explica con más detalle como hemos generado la matriz en Octave.

1. Ejemplo básico: Estabilidad y robustez

Para empezar vamos a probar con el ejemplo que utilizábamos con el Método de los Coeficientes de Fourier en el Capítulo 2 y que viene explicado en el Apéndice D. Recordamos que tenemos una señal de tamaño $N = 500$ con 20 picos ± 1 generados aleatoriamente. Gráficamente la señal era así:

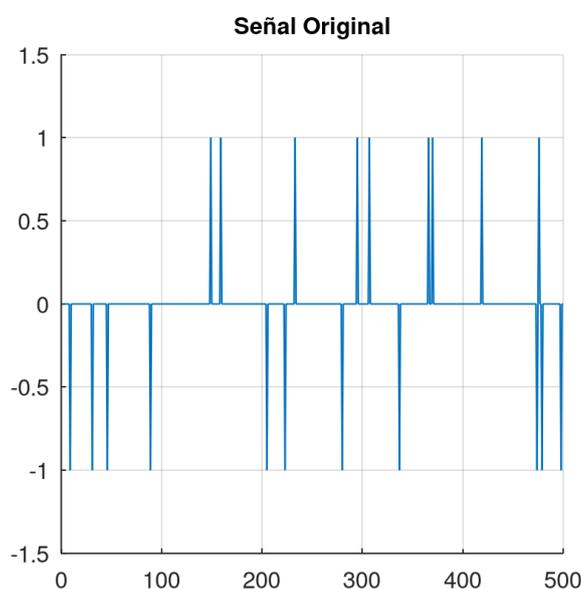


Figura 5.1: Ejemplo Básico: Señal original.

Lo primero que tenemos que determinar es el número de mediciones. Al estar utilizando matrices aleatorias y según el Teorema 3.6, el número mínimo de mediciones necesarias es $m > C \cdot s \log(N/s)$ para cierta $C > 0$ que no hemos llegado a calcular explícitamente. En este caso, con $N = 500$ y $s = 20$, tenemos

$$s \log(N/s) = 20 \log(500/20) = 64.378.$$

Como en este método entra en juego la aleatoriedad podríamos repetir la reconstrucción varias veces con la misma señal y obtener resultados diferentes. Por ejemplo, si tomamos $m = 4s = 80$ y ejecutamos el método 10 veces veremos que la reconstrucción no es siempre exacta.

Esto lo podemos ver con los errores relativos $\|x - x^*\|_2 / \|x\|_2$, donde x^* es la señal reconstruida, pues si mostramos los errores de las 10 ejecuciones obtendríamos:

```
Error =
6.8879e-01  1.2911e-04  6.9550e-01  9.0417e-01  9.0799e-01
9.5909e-01  5.5416e-01  6.7251e-05  7.8543e-01  1.6319e-04
```

Lo que nos da a entender que la reconstrucción sólo ha ido bien en 3 de las 10 ejecuciones. Gráficamente podemos mostrar la mejor y la peor reconstrucción para confirmarlo.

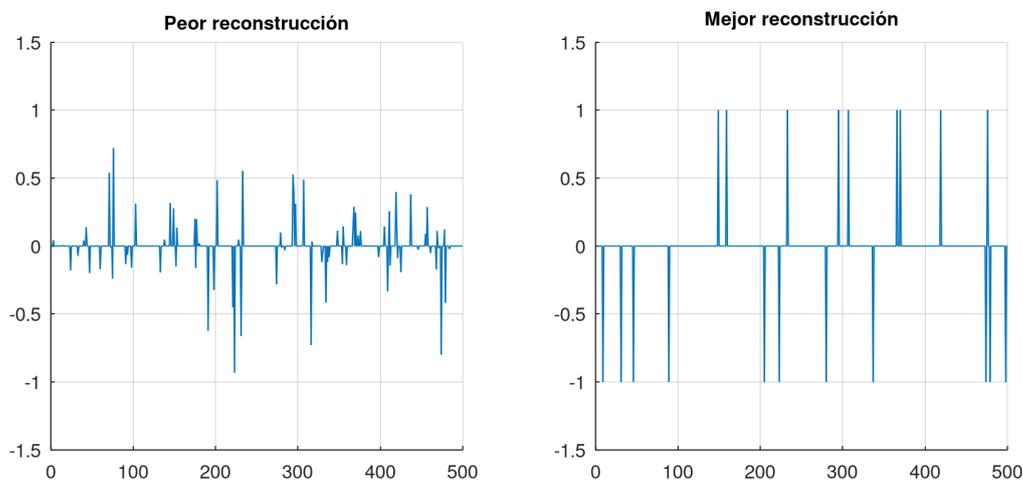


Figura 5.2: Ejemplo Básico: Mejor y peor reconstrucción con $m = 80$.

Luego, para conseguir unas reconstrucciones más fiables habrá que tomar un mayor número de mediciones. Si, por ejemplo, tomáramos ahora $m = 2 \cdot 65 = 130$, los errores de 10 reconstrucciones serían

Error =

| | | | | |
|------------|------------|------------|------------|------------|
| 9.8979e-06 | 2.6322e-06 | 4.1813e-06 | 1.0338e-05 | 1.3439e-05 |
| 2.3905e-06 | 7.3588e-06 | 2.1944e-06 | 4.1850e-06 | 1.3372e-05 |

Con ese valor de m las 10 reconstrucciones serían prácticamente exactas, lo que sugiere que la constante C puede estar cerca de 2. En lo que sigue tomaremos ese número de mediciones.

Aunque es verdad que estamos utilizando muchas más mediciones que las 40 del Método de los Coeficientes de Fourier recordamos que la Minimización ℓ_1 , como ya hemos visto en el desarrollo teórico del trabajo, será robusta frente a los errores de medición y estable si la señal no es del todo *s-sparse*. Vamos a comprobarlo con este ejemplo.

Empezaremos por la robustez añadiendo errores al vector de medición. Los errores los conseguiremos sumándole a y un vector aleatorio, E , generado por una Normal($0, \sigma^2$). En concreto, usaremos $\sigma^2 = 10^{-4}, 10^{-3}$ y 10^{-2} para ver como va empeorando la reconstrucción. Tras su generación con la Normal los errores aleatorios acaban teniendo una norma ℓ_2 igual a $\|E\|_2 = 0.1094, 0.34595$ y 1.094 respectivamente.

Si mostramos los errores relativos de la reconstrucción obtenemos:

Error =

| | | |
|----------|----------|----------|
| 0.034035 | 0.107579 | 0.340203 |
|----------|----------|----------|

Evidentemente, conforme aumenta el error de medición la reconstrucción empeora y los errores relativos aumentan. Aún así no encontramos ninguno pico extremo como si ocurría con el Método de los Coeficientes de Fourier que alcanzaba un error relativo de 4.0917 simplemente por sumar 0.001 a una componente del vector y .

Gráficamente vemos que la reconstrucción se vuelve cada vez más oscilatoria a medida que aumenta el error de medición.

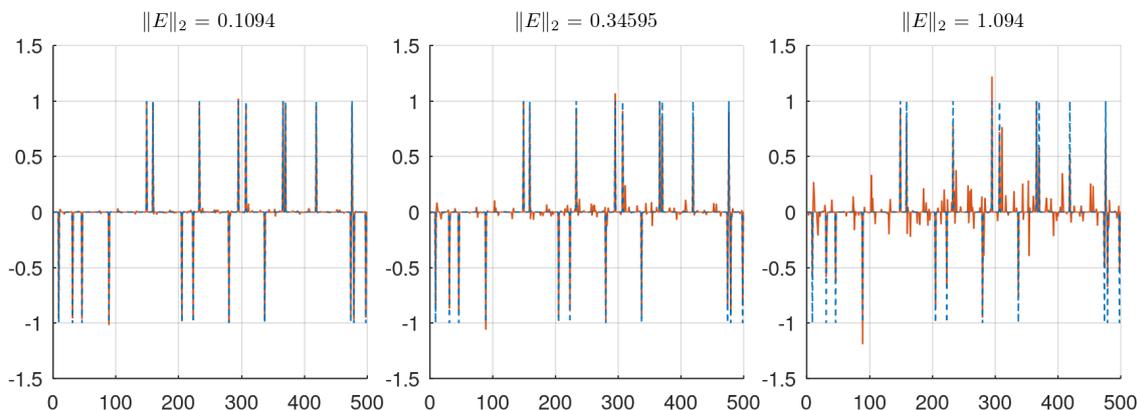


Figura 5.3: Ejemplo Básico: Robustez ante errores de medición.

Por último, para probar la estabilidad frente a que la señal no sea del todo *sparse* haremos algo similar que con las mediciones. Ahora es a la señal x a quién le sumaremos un vector aleatorio generado por una $\text{Normal}(0, 10^{-3})$.

Al hacer esto la señal dejará de ser realmente *s-sparse* pero aún así obtendremos una reconstrucción con un pequeño error relativo, 0.1882. Recordamos que con el método de los Coeficientes de Fourier simplemente con añadir 0.1 en la componente 200 ya obteníamos un error relativo de 9.7763.

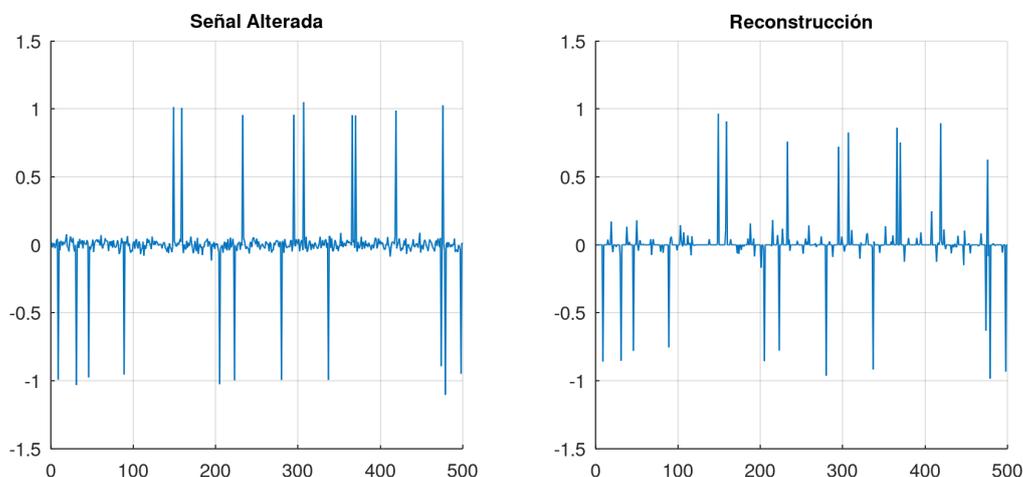


Figura 5.4: Ejemplo Básico: Estabilidad antes señales no *sparse* del todo.

2. Reconstrucción de imágenes

Ahora que hemos confirmado que la Minimización ℓ_1 funciona correctamente podemos probar con un ejemplo algo más interesante, la reconstrucción de imágenes. Al igual que en la introducción, volveremos a tomar una fotografía de la base de datos USC-SIPI Image Database [16].

Esta vez tomaremos la imagen 5.1.09 que nos muestra la superficie de la Luna. La fotografía está en escala de grises con una profundidad de 8 bits, es decir, cada píxel estará representado por un valor numérico entre 0 y 255 que nos indica su tono de gris (0 para el negro y 255 para el blanco). En cuanto a la resolución, aunque la imagen original tenía tamaño 256×256 píxeles la hemos reducido a 128×128 para que la cantidad de datos fuera más manejable.



Figura 5.5: Fotografía de la superficie de la Luna. (5.1.09)

Con todo esto, si vectorizamos la matriz de los píxeles obtendremos una señal de tamaño $N = 128^2 = 16384$. En esta ocasión no conocemos de antemano como de *sparse* es la señal por lo que habrá que calcular el valor de s . Para hacerlo vamos a expresar la señal en una base de Wavelets pues, como ya veíamos en la introducción, esta representación funciona muy bien para visualizar el nivel de *sparse* de las señales.

Aunque existen muchas bases de Wavelets que podemos utilizar, intentar explicar en que consisten, o que las diferencia unas de otras, queda fuera de nuestro alcance y requeriría un trabajo mucho más extenso. Por suerte, en Octave disponemos del paquete `l1fat` [1] y la función `fwf()` que nos calcula directamente los coeficientes de la señal en términos de la base de Wavelets que elijamos.

Para nuestro ejemplo nos bastará con utilizar la misma que proponen en el ejemplo del paquete, las Wavelets de Daubechies 8 con 10 iteraciones de filtrado, `fwf(v, 'db8', 10)`. Aún así, se puede consultar su documentación [1] para ver como esta implementado, o incluso el libro de Mallat [12] para entender como funcionan las Wavelets en general.

Si los representamos gráficamente vemos que la mayoría de coeficientes son pequeños. De hecho, si hacemos zoom vemos que sobretodo están concentrados en el intervalo $[-0.05, 0.05]$.

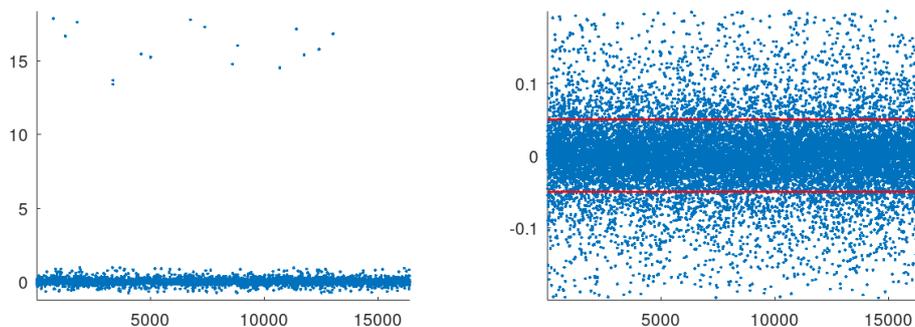


Figura 5.6: Los coeficientes de la base de Wavelets de la imagen.

Con esto en mente, podríamos ver qué le ocurriría a la imagen si hacemos cero todos los coeficientes del intervalo. Es decir, estamos considerando s como el número de coeficientes con valor absoluto mayor que 0.05. Si hacemos los cálculos obtenemos que $s = 4456$.

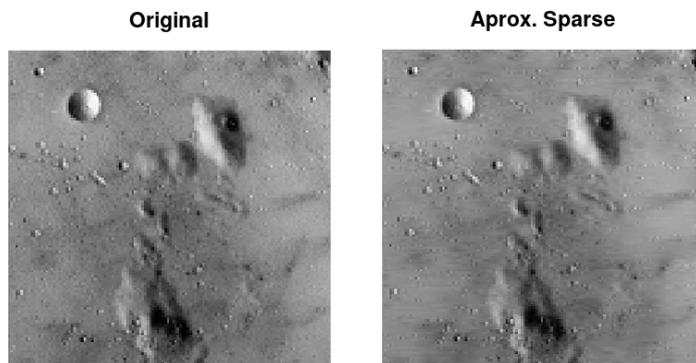


Figura 5.7: Aproximación s -sparse de la imagen.

Vemos que la aproximación s -sparse es bastante buena y que, por tanto, con tan algo menos del 30 % de los datos podemos representar fielmente la imagen. Trabajaremos con ese valor de s para la reconstrucción.

El siguiente paso es elegir el número de mediciones. Sabemos que como mínimo necesitaremos del orden de $s \log(N/s)$ mediciones que en este caso sería 5802. Podemos probar a reconstruirla con tres valores de m diferentes: $m_1 = 5802$, $m_2 = (3/2) \cdot 5802 = 8703$ y $m_3 = 2 \cdot 5802 = 11604$ para ver como va mejorando la reconstrucción.

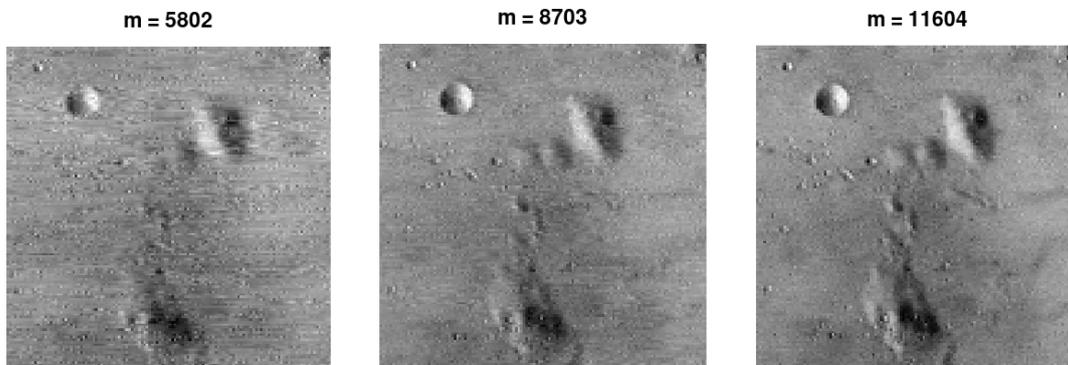


Figura 5.8: Reconstrucción mediante Minimización ℓ_1 de la imagen.

Evidentemente con este ejemplo solo pretendíamos mostrar los pasos a seguir para la reconstrucción de imágenes. Todavía se podría hacer un análisis mucho más en profundidad estudiando las diferentes bases de Wavelets o buscando otras implementaciones del método.

De todas manera, en el Apéndice D podemos encontrar el código que hemos utilizado para recrear los resultados e imágenes.

Apéndice

En este Apéndice veremos varios resultados y demostraciones que utilizamos durante el trabajo. Además, también incluiremos el código de algunos de los programas que hemos implementado en Octave [4].

A. Matrices de Vandermonde

En primer lugar, veremos las matrices de Vandermonde y daremos una fórmula para calcular su determinante. Realmente esto lo podemos encontrar en numerosas fuentes pero utilizaremos el Apéndice A.4 del libro de Foucart y Rauhut [11].

Definición A.1. Dados $x_1, x_2, \dots, x_N \in \mathbb{C}^N$ se define la matriz de Vandermonde asociada a esos valores como la matriz cuadrada $N \times N$,

$$V = V(x_1, x_2, \dots, x_N) = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{N-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \cdots & x_N^{N-1} \end{bmatrix}.$$

La propiedad más característica de estas matrices es que su determinante se puede expresar mediante la siguiente expresión.

Teorema A.1. El determinante de una matriz de Vandermonde $V = V(x_1, x_2, \dots, x_N)$ cumple

$$\det V = \prod_{1 \leq k < l \leq N} (x_l - x_k).$$

Demostración. Lo probaremos por inducción sobre N .

Para $N = 2$ está claro pues tendremos $\det V(x_1, x_2) = \begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = x_2 - x_1$.

Supongamos ahora que $N \geq 3$ y que el resultado es cierto para cualquier matriz de Vandermonde de tamaño $(N-1) \times (N-1)$. Vamos a ver que también se cumple para N .

Por las propiedades del determinante observamos que $\det V(x_1, x_2, \dots, x_N)$ es un polinomio en x_N , con grado menor o igual que $N-1$ y cuyas raíces son el resto de términos x_1, x_2, \dots, x_{N-1} . De esta forma podremos escribir

$$\det V(x_1, x_2, \dots, x_N) = c \prod_{1 \leq j < N} (x_N - x_j),$$

donde c será una constante que depende de x_1, x_2, \dots, x_{N-1} . De hecho, esa constante será el coeficiente de x_N^{N-1} en el determinante y si lo desarrollamos por la última fila llegaríamos a que $c = \det V(x_1, x_2, \dots, x_{N-1})$. Así, utilizando la hipótesis de inducción podremos sustituir el valor de c y llegar a lo que queremos:

$$\begin{aligned} \det V(x_1, x_2, \dots, x_N) &= c \prod_{1 \leq j < N} (x_N - x_j) = \\ &= \left(\prod_{1 \leq k < l \leq N-1} (x_l - x_k) \right) \left(\prod_{1 \leq j < N} (x_N - x_j) \right) = \prod_{1 \leq k < l \leq N} (x_l - x_k). \end{aligned}$$

□

Gracias a esta propiedad sabemos que si la matriz de Vandermonde está construida a partir de elementos diferentes entonces su determinante será distinto de cero.

B. Transformada de Fourier y Convolución Discretas

En esta sección, estudiaremos la Teoría de Fourier para señales discretas. La mayoría de estos resultados se pueden encontrar en el Capítulo 3 del libro de Mallat [12]. No obstante, en vez de seguirlo directamente, seguiremos el TFG de A. R. Cabuzo [6] que a su vez está basado en ese libro.

Recordamos que llamamos señal discreta de tamaño N a cualquier vector

$$f = (f(0), f(1), \dots, f(N-1)) \in \mathbb{C}^N.$$

Así, el conjunto de las señales discretas de tamaño N será el espacio vectorial \mathbb{C}^N , que consideraremos con el producto escalar

$$\langle f, g \rangle = \sum_{n=0}^{N-1} f(n) \overline{g(n)}, \quad f, g \in \mathbb{C}^N.$$

Definición B.1 (Transformada de Fourier Discreta).

Si $f = (f(0), \dots, f(N-1)) \in \mathbb{C}^N$, definimos su transformada de Fourier Discreta como la señal $\hat{f} = (\hat{f}(0), \dots, \hat{f}(N-1)) \in \mathbb{C}^N$ con coeficientes

$$\hat{f}(k) := \sum_{n=0}^{N-1} f(n) e^{-\frac{2\pi i k n}{N}}, \quad 0 \leq k < N.$$

A veces para evitar confusión denotaremos la transformada como $\mathcal{F}(f) := \hat{f}$.

Si consideramos la familia de vectores $\{e_0, e_1, \dots, e_{N-1}\}$ de \mathbb{C}^N , donde

$$e_k(n) = e^{\frac{2\pi i k n}{N}}, \quad 0 \leq n < N,$$

tendremos que $\hat{f}(k) = \langle f, e_k \rangle$ y además se cumple lo siguiente:

Teorema B.1. La familia $\{e_0, e_1, \dots, e_{N-1}\}$ es una base ortogonal de $(\mathbb{C}^N, \langle \cdot, \cdot \rangle)$.

Demostración. Como la dimensión del espacio es N , cualquier familia ortogonal de N vectores será una base ortogonal. Bastará entonces probar que $\langle e_k, e_l \rangle = N\delta_{k,l}$.

- Si $k = l$, el producto escalar nos queda

$$\langle e_l, e_l \rangle = \sum_{n=0}^{N-1} e^{\frac{2\pi i l n}{N}} e^{-\frac{2\pi i l n}{N}} = \sum_{n=0}^{N-1} 1 = N.$$

- Si $k \neq l$, usando la fórmula de las series geométricas tendremos

$$\langle e_k, e_l \rangle = \sum_{n=0}^{N-1} e^{\frac{2\pi i (k-l)n}{N}} = \frac{1 - \left(e^{\frac{2\pi i (k-l)}{N}} \right)^N}{1 - e^{\frac{2\pi i (k-l)}{N}}} = \frac{1 - e^{2\pi i (k-l)}}{1 - e^{\frac{2\pi i (k-l)}{N}}} = \frac{1 - 1}{1 - e^{\frac{2\pi i (k-l)}{N}}} = 0.$$

Observamos que el denominador no se anula ya que $k \neq l$ y $|k - l| < N$.

□

Ahora que ya sabemos que es una base ortogonal podremos probar la llamada Fórmula de Inversión de la transformada de Fourier discreta.

Teorema B.2 (Fórmula de Inversión).

Sea $f \in \mathbb{C}^N$ una señal discreta, entonces $f(n) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}(k) e^{\frac{2\pi i k n}{N}}$, $0 \leq n < N$.

Demostración. Como la familia $\{e_k\}$ forma una base ortogonal, cualquier señal $f \in \mathbb{C}^N$ se puede escribir como

$$f = \sum_{n=0}^{N-1} \lambda_k e_k.$$

Podemos calcular λ_k haciendo

$$\langle f, e_k \rangle = \left\langle \sum_{n=0}^{N-1} \lambda_n e_n, e_k \right\rangle = \sum_{n=0}^{N-1} \lambda_n \langle e_n, e_k \rangle = \lambda_k \langle e_k, e_k \rangle.$$

Luego, $\lambda_k = \frac{\langle f, e_k \rangle}{\langle e_k, e_k \rangle}$ y como teníamos que $\langle f, e_k \rangle = \hat{f}(k)$ y $\langle e_k, e_k \rangle = N$, nos quedará

$$f = \frac{1}{N} \sum_{n=0}^{N-1} \hat{f}(k) e_k.$$

Así, su componente n -ésima será $f(n) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}(k) e^{\frac{2\pi i k n}{N}}$, $0 \leq n < N$.

□

Siguiendo este resultado podemos definir la transformada inversa.

Definición B.2 (Transformada de Fourier Inversa).

Si $g \in \mathbb{C}^N$, definimos su transformada de Fourier Inversa como la señal $\check{g} \in \mathbb{C}^N$ con coeficientes

$$\check{g}(n) = \frac{1}{N} \sum_{k=0}^{N-1} g(k) e^{\frac{2\pi i k n}{N}}, \quad 0 \leq n < N.$$

Al igual que antes, a veces denotaremos la transformada inversa como $\mathcal{F}^{-1}(g) = \check{g}$.

Por supuesto, como consecuencia del Teorema B.2 tendremos que $\mathcal{F}^{-1}(\mathcal{F}(g)) = g$, para todo $g \in \mathbb{C}^N$. También es sencillo comprobar que $\mathcal{F}(\mathcal{F}^{-1}(g)) = g, \forall g \in \mathbb{C}^N$.

El siguiente paso será definir la convolución discreta de señales. Dadas $f, h \in \mathbb{C}^N$ nos gustaría definir su convolución como

$$f * h(n) = \sum_{m=0}^{N-1} f(m)h(n-m).$$

Pero observamos que así será necesario tener valores fuera de $0 \leq n < N$. Para solucionarlo extenderemos las señales $f \in \mathbb{C}^N$ como sucesiones N -periódicas con índices en \mathbb{Z} , es decir,

$$f(n) = f(n \text{ mód } N), n \in \mathbb{Z}.$$

Por ejemplo, la familia $\{e_k\}$ es toda N -periódica ya que,

$$e_k(n+N) = e^{\frac{2\pi ik(n+N)}{N}} = e^{\frac{2\pi ikn}{N}} e^{\frac{2\pi ikN}{N}} = e^{\frac{2\pi ikn}{N}} e^{2\pi ik} = e^{\frac{2\pi ikn}{N}} = e_k(n), \forall n \in \mathbb{Z}.$$

Definición B.3 (Convolución Discreta).

Dadas $f, h \in \mathbb{C}^N$, se define su convolución discreta como la señal $f * h \in \mathbb{C}^N$ dada por

$$f * h(n) := \sum_{m=0}^{N-1} f(m)h(n-m), 0 \leq n < N.$$

Observamos que al considerar f y h como señales N -periódicas su convolución también lo será. Además, es fácil ver que $f * h = h * f$.

Con los siguientes resultados vamos a ver cómo se relacionan la convolución y la transformada de Fourier discretas.

Teorema B.3. Si $f, h \in \mathbb{C}^N$, entonces su convolución $f * h$ cumple $\widehat{f * h} = \widehat{f} \cdot \widehat{h}$, es decir,

$$\widehat{f * h}(k) = \widehat{f}(k)\widehat{h}(k), 0 \leq k < N.$$

Demostración. Por la definición de transformada y convolución tenemos

$$\begin{aligned} \widehat{f * h}(k) &= \sum_{n=0}^{N-1} (f * h)(n) e^{-\frac{2\pi ikn}{N}} = \sum_{n=0}^{N-1} \left[\sum_{p=0}^{N-1} f(p)h(n-p) \right] e^{-\frac{2\pi ikn}{N}} = \\ &= \sum_{n=0}^{N-1} \left[\sum_{p=0}^{N-1} f(p)h(n-p) \right] e^{-\frac{2\pi ik(n-p)}{N}} e^{-\frac{2\pi ikp}{N}} = \\ &= \sum_{p=0}^{N-1} \left[\sum_{n=0}^{N-1} h(n-p) e^{-\frac{2\pi ik(n-p)}{N}} \right] f(p) e^{-\frac{2\pi ikp}{N}}. \end{aligned}$$

Si en la suma de dentro del paréntesis hacemos el cambio de variable $m = n - p$ obtenemos

$$\begin{aligned} \sum_{n=0}^{N-1} h(n-p) e^{-\frac{2\pi ik(n-p)}{N}} &= \sum_{m=-p}^{N-1-p} h(m) e^{-\frac{2\pi ikm}{N}} = \\ &= \sum_{m=-p}^{-1} h(m) e^{-\frac{2\pi ikm}{N}} + \sum_{m=0}^{N-1-p} h(m) e^{-\frac{2\pi ikm}{N}} = \\ &= \sum_{m=-p}^{-1} h(m+N) e^{-\frac{2\pi ik(m+N)}{N}} + \sum_{m=0}^{N-1-p} h(m) e^{-\frac{2\pi ikm}{N}}. \end{aligned}$$

En el último paso hemos utilizado la N -periodicidad de las señales. Si volvemos a hacer un cambio de variable, ahora $l = m + N$, en el primer sumando tendremos

$$\begin{aligned} \sum_{n=0}^{N-1} h(n-p)e^{-\frac{2\pi ik(n-p)}{N}} &= \sum_{l=N-p}^{N-1} h(l)e^{-\frac{2\pi ikl}{N}} + \sum_{m=0}^{N-1-p} h(m)e^{-\frac{2\pi ikm}{N}} = \\ &= \sum_{l=0}^{N-1} h(l)e^{-\frac{2\pi ikl}{N}} = \widehat{h}(k). \end{aligned}$$

Sustituyendo donde antes obtendremos lo que queríamos.

$$\widehat{f * h}(k) = \sum_{p=0}^{N-1} \widehat{h}(k) f(p) e^{-\frac{2\pi ikp}{N}} = \widehat{h}(k) \sum_{p=0}^{N-1} f(p) e^{-\frac{2\pi ikp}{N}} = \widehat{h}(k) \widehat{f}(k) = \widehat{f}(k) \widehat{h}(k).$$

□

Para la transformada inversa y con un desarrollo similar se puede probar:

Teorema B.4. Si $f, h \in \mathbb{C}^N$, entonces su convolución $f * h$ cumple $\widetilde{f * h} = N(\check{f} \cdot \check{h})$, es decir,

$$\widetilde{f * h}(k) = N\check{f}(k)\check{h}(k), \quad 0 \leq k < N.$$

Demostración. La prueba es idéntica a la del teorema anterior. Solo hay que tener en cuenta el $1/N$ y el cambio del signo de la exponencial en la transformada inversa. Así, por definición, tendremos

$$\widetilde{f * h}(k) = \frac{1}{N} \sum_{n=0}^{N-1} (f * h)(n) e^{\frac{2\pi ikn}{N}} = \dots = \frac{1}{N} \sum_{p=0}^{N-1} \left[\sum_{n=0}^{N-1} h(n-p) e^{\frac{2\pi ikn(n-p)}{N}} \right] f(p) e^{\frac{2\pi ikp}{N}}.$$

Si en la suma de dentro del paréntesis hacemos los mismos cambios de variables y cálculos que antes obtendremos

$$\sum_{n=0}^{N-1} h(n-p) e^{\frac{2\pi ikn(n-p)}{N}} = \dots = \sum_{l=0}^{N-1} h(l) e^{\frac{2\pi ikl}{N}} = N\check{h}(k).$$

Luego, nos queda

$$\widetilde{f * h}(k) = \frac{1}{N} \sum_{p=0}^{N-1} N\check{h}(k) f(p) e^{\frac{2\pi ikp}{N}} = \check{h}(k) \sum_{p=0}^{N-1} f(p) e^{\frac{2\pi ikp}{N}} = \check{h}(k) N\check{f}(k) = N\check{f}(k)\check{h}(k).$$

□

Para acabar vamos a ver qué ocurre cuando en vez de hacer la convolución de señales hacemos la convolución de sus transformadas o transformadas inversas.

Teorema B.5. Sean $F, H \in \mathbb{C}^N$, la convolución de sus transformadas inversas cumple

$$\check{F} * \check{H} = \widetilde{F \cdot H}.$$

Demostración. Llamamos $f = \check{F} \in \mathbb{C}^N$ y $h = \check{H} \in \mathbb{C}^N$, de manera que, $\hat{f} = F$ y $\hat{h} = H$.

Entonces por el Teorema B.3 tendremos $\widehat{f * h} = \hat{f} \cdot \hat{h}$ y podremos escribir

$$\mathcal{F}(\check{F} * \check{H}) = \widehat{f * h} = \hat{f} \cdot \hat{h} = F \cdot H.$$

Basta tomar la transformada inversa y llegamos a lo que queríamos

$$\check{F} * \check{H} = \widetilde{F \cdot H}.$$

□

Teorema B.6. Sean $f, h \in \mathbb{C}^N$, la convolución de sus transformadas cumple

$$\hat{f} * \hat{h} = N \widehat{f \cdot h}.$$

Demostración. Llamamos $F = \hat{f} \in \mathbb{C}^N$ y $H = \hat{h} \in \mathbb{C}^N$, de manera que, $\check{F} = f$ y $\check{H} = h$.

Entonces por el Teorema B.4 tendremos $\widetilde{F * H} = N(\check{F} \cdot \check{H})$ y podremos escribir

$$\mathcal{F}^{-1}(\hat{f} * \hat{h}) = \widetilde{F * H} = N(\check{F} \cdot \check{H}) = N(f \cdot h).$$

Basta tomar la transformada y llegamos a lo que queríamos

$$\hat{f} * \hat{h} = N \widehat{f \cdot h}.$$

□

En definitiva, para dos señales $f, h \in \mathbb{C}^N$, hemos probado que:

| | |
|--|---|
| $\widehat{f * h} = \hat{f} \cdot \hat{h}$ | $\hat{f} * \hat{h} = N \widehat{f \cdot h}$ |
| $\widetilde{f * h} = N(\check{f} \cdot \check{h})$ | $\check{f} * \check{h} = \widetilde{f \cdot h}$ |

C. Conceptos Básicos de Probabilidad

En esta sección repasaremos algunos conceptos básicos de Teoría de la Probabilidad. La mayoría de estos los podemos encontrar en el Capítulo 7 del libro [11], aunque también consultaremos el manual [13].

Tenemos un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, donde Ω es el espacio muestral, \mathcal{F} es una σ -álgebra de subconjuntos de Ω y \mathbb{P} es una medida de probabilidad. Llamaremos suceso a cualquier elemento $A \in \mathcal{F}$ y denotaremos su probabilidad por

$$\mathbb{P}(A) = \int_A d\mathbb{P}(\omega) = \int_{\Omega} I_A(\omega) d\mathbb{P}(\omega),$$

con $I_A(\cdot)$ la función característica de A que vale 1 si $\omega \in A$ y 0 en otro caso.

Para una familia finita de sucesos, $\{A_i\}_{i=1}^n \subseteq \mathcal{F}$, la probabilidad de su unión está acotada por la suma de sus probabilidades, es decir, se cumple la desigualdad

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i).$$

Definición C.1 (Variable aleatoria).

Llamaremos *variable aleatoria* a cualquier función $X : \Omega \rightarrow \mathbb{R}$ medible en (Ω, \mathcal{F}) . Recordamos que para ser medible, X debe cumplir que

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$$

para cualquier B perteneciente a la σ -álgebra de Borel de \mathbb{R} , \mathcal{B} .

La variable X junto con la medida \mathbb{P} nos permiten definir una medida de probabilidad, \mathbb{P}_X , en $(\mathbb{R}, \mathcal{B})$ haciendo

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)).$$

Además, podemos definir su *función de distribución*, F_X , como

$$F_X(t) = \mathbb{P}_X((-\infty, t]) = \mathbb{P}(X \leq t).$$

Se puede ver que la función de distribución nos permite determinar la medida \mathbb{P}_X .

Definición C.2 (Idénticamente distribuidas).

Si $\{X_i\}_{i \in I}$ es una familia de variables aleatorias tal que $\mathbb{P}_{X_i} = \mathbb{P}_{X_j}$, $\forall i, j \in I$ entonces diremos que las variables son *idénticamente distribuidas*. Como las funciones de distribución nos determinan las medidas podríamos también decir lo mismo si $F_{X_i} = F_{X_j}$, $\forall i, j \in I$.

Para cualquier colección finita de variables aleatorias X_1, \dots, X_n podemos considerar (X_1, \dots, X_n) como una función de Ω a \mathbb{R}^n que nos determina una medida $\mathbb{P}_{(X_1, \dots, X_n)}$ sobre \mathbb{R}^n . A esta se le conoce como la *distribución conjunta* de (X_1, \dots, X_n) .

Definición C.3 (Independientes).

Diremos que una familia de variables aleatorias, $\{X_i\}_{i \in I}$, son *independientes* si la distribución conjunta de cualquier colección finita de variables de la familia es el producto de las distribuciones individuales de las variables. Es decir, si (X_1, \dots, X_n) pertenecen a $\{X_i\}_{i \in I}$ entonces se cumple

$$\mathbb{P}_{(X_1, \dots, X_n)} = \prod_{i=1}^n \mathbb{P}_{X_i}.$$

Cuando una familia $\{X_i\}_{i \in I}$ cumple ambas definiciones diremos que son *independientes e idénticamente distribuidas* y lo abreviaremos por i.i.d.

Definición C.4 (Esperanza).

Se define la *esperanza* o valor esperado de la variable aleatoria X como

$$\mathbb{E}[X] := \int_{\Omega} X(\omega) d\mathbb{P}(\omega).$$

Por la definición es sencillo ver que la esperanza es un operador lineal, es decir, para cualesquiera variables aleatorias X e Y y cualquier $c \in \mathbb{R}$, se cumple

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] \quad \text{y} \quad \mathbb{E}[cX] = c\mathbb{E}[X].$$

Además, si X está delimitada por dos números reales, $a < X < b$, entonces también lo está su esperanza, $a < \mathbb{E}[X] < b$, y si $X \leq Y$ entonces $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

Una de las propiedades que más utilizaremos es que para una familia $\{X_i\}_{i=1}^n$ de variables aleatorias independientes se cumple

$$\mathbb{E}\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n \mathbb{E}(X_i).$$

También es cierto el siguiente lema:

Lema C.1. Sea $H : [0, \infty) \rightarrow [0, \infty)$ derivable y creciente con $H(0) = 0$. Entonces para toda variable aleatoria X se tiene

$$\mathbb{E}[H(|X|)] = \int_0^\infty H'(t) \mathbb{P}(|X| > t) dt.$$

Demostración. Usando la regla de Barrow en la segunda línea y el teorema de Fubini en la tercera, se obtiene

$$\begin{aligned} \mathbb{E}[H(|X|)] &= \int_{\Omega} H(|X(w)|) d\mathbb{P}(w) = \\ &= \int_{\Omega} \left[\int_0^{|X(w)|} H'(t) dt \right] d\mathbb{P}(w) = \\ &= \int_0^\infty H'(t) \int_{\{w \in \Omega : |X(w)| > t\}} d\mathbb{P}(w) dt = \\ &= \int_0^\infty H'(t) \mathbb{P}(|X| > t) dt. \end{aligned}$$

□

Distinguiremos dos tipos de variable aleatoria en función de como sea su imagen, $X(\Omega)$.

Caso Discreto:

La imagen es un conjunto finito o numerable de puntos en \mathbb{R} . De esta manera, para $B \in \mathcal{B}$, tendremos

$$\mathbb{P}_X(B) = \sum_{\{x_i \in X(\Omega) : x_i \in B\}} \mathbb{P}(X = x_i).$$

Evidentemente, al ser una medida de probabilidad se debe cumplir

$$\sum_{\{x_i \in X(\Omega)\}} \mathbb{P}(X = x_i) = \mathbb{P}_X(\mathbb{R}) = 1.$$

En este caso, podremos calcular la esperanza de la variable discreta X como

$$\mathbb{E}[X] = \sum_{\{x_i \in X(\Omega)\}} x_i \mathbb{P}(X = x_i).$$

Caso Continuo:

La imagen es un conjunto no numerable de \mathbb{R} , es decir, $X(\Omega)$ abarcará todo un intervalo de puntos. En este caso, si además \mathbb{P}_X es absolutamente continua respecto a la medida de Lebesgue, entonces existirá una función medible $f_X : \mathbb{R} \rightarrow [0, +\infty)$ tal que

$$\mathbb{P}_X([a, b]) = \mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x) dx, \quad \text{para todo } a < b.$$

Se conoce a f_X como la *función de densidad* de la variable aleatoria X y se relaciona con la función de distribución pues

$$F_X(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t f_X(x) dx.$$

De hecho, se cumple que $f_X(x) = \frac{d}{dx} F_X(x)$. Evidentemente, al estar en un espacio de probabilidad,

$$\int_{-\infty}^{+\infty} f_X(x) dx = \mathbb{P}_X(\mathbb{R}) = 1.$$

Además, podremos calcular la esperanza de la variable continua X como

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x) dx.$$

A parte de la esperanza podemos definir otras medidas que sintetizan el comportamiento de las variables aleatorias.

Definición C.5 (Momentos).

Se define el *momento de orden k* de una variable aleatoria X como

$$\mathbb{E}[X^k], \text{ siempre que } \mathbb{E}[|X|^k] < \infty.$$

Observamos que el momento de orden 1 no es más que la esperanza de la variable.

Definición C.6 (Función generatriz de momentos).

La *función generatriz de momentos* de una variable aleatoria X es la función

$$M_X(t) := \mathbb{E}[e^{tX}], \text{ para todo } t \in \mathbb{R}.$$

Como $e^{tX} > 0$, la función $M(t)$ está siempre bien definida, si admitimos que toma valores en $[0, +\infty]$. Se le llama función generatriz de momentos porque nos permite generar los momentos de la distribución,

$$\mathbb{E}[X^k] = \frac{d^k M_X}{dt^k}(0).$$

D. Códigos de Octave

Para acabar el Apéndice incluiremos los códigos de varios programas que hemos desarrollado en el trabajo.

Método de los coeficientes de Fourier

El primer programa que incluimos es la implementación del método de reconstrucción a partir de los primeros $2s$ coeficientes de la transformada de Fourier de la señal que estudiamos en el Teorema 2.3.

Para la implementación del programa y su posterior uso con algunos ejemplos, fueron necesarias las funciones `fft()` y `ifft()` que nos permiten calcular la transformada de Fourier discreta y su inversa en Octave [4].

```

% Parámetros:
% N = Tamaño de la señal.
% s = Nº de coeficientes no nulos de la señal (nivel de sparse).
% y = Los 2s primeros coeficientes de la transformada de Fourier
%     de la señal.

% x0 = La reconstrucción de la señal obtenida.

function x0 = FourierMethod(N,s,y)

    % Calculamos las componentes 1,...,s de la transformada de q
    A = zeros(s,s);
    b = zeros(s,1);
    for j=1:s
        for k=1:s
            A(j,k) = y(s+j-k);
        endfor
        b(j)=-y(s+j);
    endfor
    coef = A\b;

    % El resto de la transformada era TFq(0)=1 y TFq(k)=0 si k > s
    TFq = zeros(N,1);
    TFq(1) = 1;
    for j=1:s
        TFq(j+1) = coef(j);
    endfor
    q = ifft(TFq);    % Calculamos la inversa para obtener q

    % Nos quedamos con los índices de los s coeficientes más pequeños
    % que serán las raíces de q, es decir, el soporte T de la señal
    T = find(abs(q)<=abs(sort(q)(s)));

    % Calculamos los coeficientes x(t) en T que son los únicos no nulos
    A = zeros(2*s,s);
    for j=1:2*s
        for k=1:s
            A(j,k) = exp(-2*pi*sqrt(-1)*(j-1)*(T(k)-1)/N);
        endfor
    endfor
    z = A\y;

    % El resto de la señal será toda ceros
    x0 = zeros(N,1);
    for j=1:s
        x0(T(j))=z(j);
    endfor

endfunction

```

Ejemplo básico

Ahora mostraremos como hemos generado la señal de tamaño $N = 500$ compuesta por $s = 20$ picos ± 1 que utilizamos con los métodos ℓ_0 y ℓ_1 para comprobar su eficacia, estabilidad y robustez. Este código es el mismo que se utiliza en los ejemplos del paquete ℓ_1 -MAGIC [8].

```

rng(2);
N = 500;
s = 20;

```

```
% Señal +/- 1 aleatoria
x = zeros(N,1);
perm = randperm(N);
x(perm(1:s)) = sign(randn(s,1));
```

La única diferencia es que añadimos el comando `rng(2)` para fijar la semilla de generación de números aleatorios y así obtener la misma señal todas las veces.

Para la reconstrucción por el Método de los Coeficientes de Fourier las mediciones serán los primeros $2s$ coeficientes de la transformada discreta de la señal. Usando la función `fft()` que calcula una Transformada Rápida de Fourier en Octave haríamos:

```
y = fft(x);
y = y(1:2*s);

% Y llamamos al algoritmo para la reconstrucción
x0 = FourierMethod(N,s,y);
```

Para la reconstrucción mediante la Minimización en norma ℓ_1 usaremos una matriz de mediciones aleatoria. En concreto, como queremos una distribución Estrictamente Subgaussiana usaremos la Normal($0, 1/m$). Para construirla usaremos el paquete `statistics` y la función `normrnd(mu, sigma, m, N)` que genera una matriz de tamaño $m \times N$ a partir de una Normal(μ, σ^2). El código sería:

```
rng shuffle
A = normrnd(0, 1/sqrt(m), m, N);

y = A*x;           # Las mediciones
x0 = A'*y;         # Posible punto inicial para el algoritmo

% Y llamamos al algoritmo para la reconstrucción
xp = l1eq_pd(x0, A, [], y);
```

Como este procedimiento se basa en una construcción aleatoria hemos añadido el comando `rng shuffle` para resetear la semilla de la generación de números aleatorios. Sino aún seguiríamos en la semilla `rng(2)` y siempre obtendríamos la misma matriz de mediciones.

Para comprobar la robustez de los métodos hemos tenido que generar errores en la medición. En el método de los coeficientes nos ha bastado con hacer $y(s) = y(s) + 0.001$ pero para el método ℓ_1 hemos vuelto a utilizar la función `normrnd()` para añadir error en todas las mediciones haciendo $E = \text{normrnd}(0, 1e-2, m, 1)$ y luego $y = y + E$.

De la misma manera, para comprobar la estabilidad modificábamos la señal original x para que no fuera del todo s -sparse. En los coeficientes de Fourier simplemente le dábamos el valor 0.1 a la componente que queríamos, $x(t) = 0.1$, y en la Minimización ℓ_1 hemos hecho $x = x + \text{normrnd}(0, \text{sqrt}(10^{-3}), N, 1)$.

Reconstrucción de imágenes

Para acabar mostraremos el código utilizado para la reconstrucción de la imagen 5.1.09 de la base de datos [16]. En primer lugar leemos los datos, convertimos la matriz de píxeles, con valores enteros entre 0 y 255, a una matriz real con valores entre 0 y 1 para poder operar con ella. Luego, vectorizamos la matriz para obtener la señal y la expresamos en términos de una base de Wavelets con la función `fwf()` del paquete [1].

```

# Leemos la imagen y la mostramos por pantalla
img = imread ("Luna.tif");

figure
imshow(img)
title("Original")

% Obtenemos la longitud de la señal
d = size(img)(1);      # En este caso d = 128
N = d^2;              # N = 128^2 = 16384

% Transformamos la matriz de píxeles a real y la vectorizamos
I = im2double(img);
v = I'(:);

% Cargamos el paquete ltfat y expresamos la señal en la base de Wavelets
pkg load ltfat
x = fwt(v, 'db8', 10);

```

Para descubrir el nivel de *sparse* de la señal podemos representar gráficamente los coeficientes. Al final decíamos que s sería el número de coeficientes con valor absoluto mayor que 0.05. Una vez obtenido s mostrábamos la aproximación s -*sparse* de la imagen.

```

# Mostramos los coeficientes permutados para mejorar su visibilidad
figure
hold on
perm = randperm(N);
plot(x(perm(1:N)), '.');
axis([1 N min(x)-0.5 max(x)+0.5])

# Obtenemos el nivel de sparse
s = sum(abs(x)>0.05)

# Hacemos 0 los coeficientes con valor absoluto menor que 0.05
xs = x.*(abs(x)>0.05);

# Calculamos la inversa de la transformada de Wavelets
Is = ifwt(xs, 'db8', 10, N);
# Nos aseguramos que solo tiene valores entre 0 y 1
Is = rescale(Is, 0, 1);
# Pasamos de un vector a una matriz 128x128
Is = reshape(Is, d, d)';
# Transformamos la matriz de reales en [0,1] en una
# imagen en escala de grises con profundidad 8 bits
Is = gray2ind(Is, 256);

# Mostramos la aproximación sparse
figure
imshow(Is)
title("Aprox. Sparse")

```

Por último calculamos el número mínimo de mediciones necesarias y con él hacemos la reconstrucción con la Minimización en norma ℓ_1 . Por supuesto, una vez tenemos la señal recuperada construimos con ella una imagen y la mostramos por pantalla.

```

# Número de mediciones
m = ceil(s*log(N/s))

# Generamos la matriz de mediciones
pkg load statistics
A = normrnd(0, 1/sqrt(m), m, N);

```

```
# Obtenemos las mediciones y el punto inicial para el algoritmo
y = A*x;
x0 = A'*y;

# Llamamos al método de reconstrucción
xp = l1eq_pd(x0, A, [], y);

# Mostramos la imagen reconstruida
Ip = ifwt(xp, 'db8', 10, N);
Ip = rescale(Ip, 0, 1);
Ip = reshape(Ip, d, d)';
Ip = gray2ind(Ip, 256);

figure
imshow(Ip)
title("Reconstrucción")
```


Referencias

- [1] P. Balazs, P. L. Søndergaard y B. Torrèsani. *The Linear Time Frequency Analysis Toolbox*. International Journal of Wavelets, Multiresolution and Information Processing. **10**(2012), núm 4.
Disponible en: <http://lftfat.org/>
- [2] R. Baraniuk, M. A. Davenport, R. DeVore y M. Wakin. *A Simple Proof of the Restricted Isometry Property for Random Matrices*. Constructive Approximation. **28**(2008), 253-268.
- [3] R. Baraniuk, M. A. Davenport, M. F. Duarte y C. Hegde. *An Introduction to Compressive Sensing*. 2011.
Disponible en: <https://legacy.cnx.org/content/col111133/1.5/>
- [4] D. Bateman, J. W. Eaton, S. Hauberg y R. Wehbring. *GNU Octave version 7.1.0 manual: a high-level interactive language for numerical computations*. 2022.
Disponible en: <https://www.gnu.org/software/octave/doc/v7.1.0/>
- [5] S. Boyd y L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [6] A. R. Cabuzo. *La transformada de Fourier discreta y el formato JPEG*. TFG, Universidad de Murcia, 2018. Disponible en:
https://webs.um.es/gustavo.garrigos/tfg/Cabuzo_TFG_enero2018.pdf
- [7] E. J. Candès. *The restricted isometry property and its implications for compressed sensing*. Comptes Rendus Mathématique. **346**(2008), núm. 9, 589-592.
- [8] E. J. Candès y J. Romberg. *ℓ_1 -MAGIC: Recovery of Sparse Signals via Convex Programming*. Caltech, 2005.
Disponible en: <https://candes.su.domains/software/l1magic/>
- [9] B. Cascales, J.M. Mira, J. Orihuela y M. Raja. *Análisis Funcional*. Electolibris, 2012.
- [10] R. DeVore, G. Petrova y P. Wojtaszczyk. *Instance-optimality in probability with an ℓ_1 -minimization decoder*. Appl. Comput. Harmon. Anal. **28**(2009), núm. 3, 275-288.
- [11] S. Foucart y H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Springer New York, 2013.
- [12] S. Mallat. *A Wavelet Tour of Signal Processing (3rd Ed.)*. Elsevier, 2009.
- [13] F. Montes. *Introducción a la Probabilidad*. Universitat de València, 2007.
Disponible en: <https://www.uv.es/montes/probabilitat/manual.pdf>
- [14] S. Shalev-Shwartz. *Compressed Sensing: Basic results and self contained proofs*. 2009.
Disponible en: <https://www.cs.huji.ac.il/~shais/compressedSensing.pdf>

- [15] O. Rivasplata. *Subgaussian random variables: An expository note*. 2012.
Disponibile en: <https://www.stat.cmu.edu/~arinaldo/36788/subgaussians.pdf>
- [16] *The USC-SIPI Image Database*. Signal and Image Processing Institute, University of Southern California, 2018.
Disponibile en: <https://sipi.usc.edu/database/database.php>