Adquisición Compresiva (Compressed Sensing)

Carlos Fernández Granda

Departamento de Ingeniería Eléctrica
Universidad de Stanford

19/7/2011

Índice

Imagen por Resonancia Magnética

- **Modelo** : Medimos la transformada de Fourier de una sección del cuerpo del paciente.

Imagen por Resonancia Magnética

- **Modelo** : Medimos la transformada de Fourier de una sección del cuerpo del paciente.
- **Problema** : Duración prolongada del proceso de medida.

Imagen por Resonancia Magnética

- **Modelo** : Medimos la transformada de Fourier de una sección del cuerpo del paciente.
- **Problema** : Duración prolongada del proceso de medida.
  - Incómodo para el paciente.

Imagen por Resonancia Magnética

- **Modelo** : Medimos la transformada de Fourier de una sección del cuerpo del paciente.
- **Problema** : Duración prolongada del proceso de medida.
  - Incómodo para el paciente.
  - Limita ciertas aplicaciones (imágenes cardiacas).

Imagen por Resonancia Magnética

- **Modelo** : Medimos la transformada de Fourier de una sección del cuerpo del paciente.
- **Problema** : Duración prolongada del proceso de medida.
    - Incómodo para el paciente.
    - Limita ciertas aplicaciones (imágenes cardiacas).
- Sin embargo, con frecuencia la imagen es **compresible**.

Compresión mediante Wavelets

- **Transformada wavelet** : Descomposición en una base de funciones con distintos niveles de resolución.

Compresión mediante Wavelets

- **Transformada wavelet** : Descomposición en una base de funciones con distintos niveles de resolución.
- Muchas imágenes naturales son dispersas (*sparse*) en estas bases.

Compresión mediante Wavelets

- **Transformada wavelet** : Descomposición en una base de funciones con distintos niveles de resolución.
- Muchas imágenes naturales son dispersas (*sparse*) en estas bases.

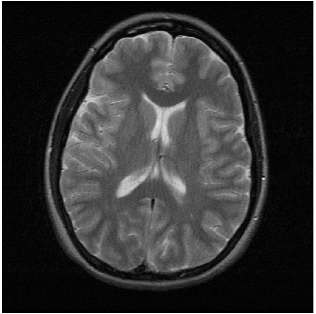
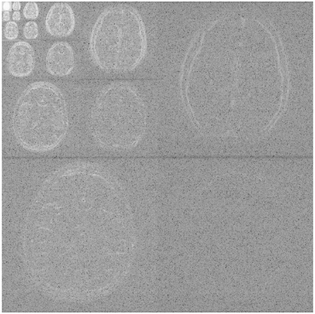


Imagen original



Coeficientes Wavelet

Compresión mediante Wavelets

Si eliminamos un 90% de los coeficientes de la descomposición, la calidad apenas se resiente.

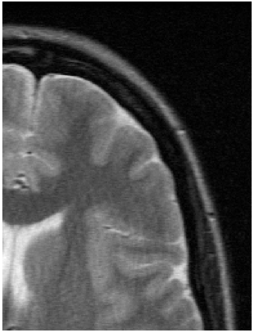


Imagen original

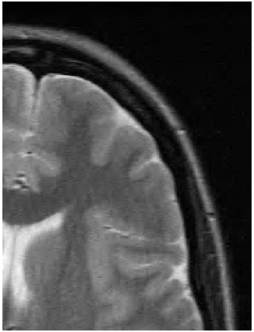


Imagen comprimida

Aliasing

- Estamos midiendo información muy redundante.

Aliasing

- Estamos midiendo información muy redundante.
- ¿Por qué no submuestrear?

Aliasing

- Estamos midiendo información muy redundante.
- ¿Por qué no submuestrear ?
- Porque lo dice Nyquist (y Shannon).

Aliasing

- Estamos midiendo información muy redundante.
- ¿Por qué no submuestrear?
- Porque lo dice Nyquist (y Shannon).
- El submuestreo produce solapamientos (*aliasing*).

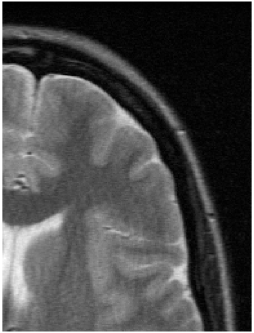
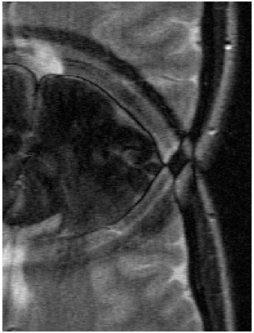


Imagen original



Submuestreo 50%

Adquisición compresiva

- **Paso 1** : Aleatorizamos el submuestreo.

Adquisición compresiva

- **Paso 1** : Aleatorizamos el submuestreo.
- **Paso 2** : Al reconstruir penalizamos la norma ℓ_1 de la transformada wavelet de la imagen.

Adquisición compresiva

- **Paso 1** : Aleatorizamos el submuestreo.
- **Paso 2** : Al reconstruir penalizamos la norma ℓ_1 de la transformada wavelet de la imagen.

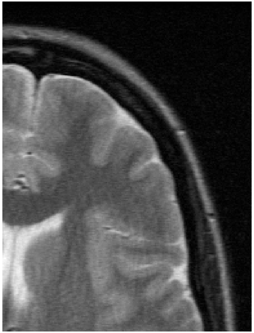


Imagen original

Submuestreo 33%

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.
- En aplicaciones modernas como tratamiento de imágenes o vídeo, las dimensiones de las señales son gigantescas.

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.
- En aplicaciones modernas como tratamiento de imágenes o vídeo, las dimensiones de las señales son gigantescas.
- Invertimos mucho esfuerzo en medir información redundante para luego desecharla.

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.
- En aplicaciones modernas como tratamiento de imágenes o vídeo, las dimensiones de las señales son gigantescas.
- Invertimos mucho esfuerzo en medir información redundante para luego desecharla.
- Paradigma de la adquisición compresiva :

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.
- En aplicaciones modernas como tratamiento de imágenes o vídeo, las dimensiones de las señales son gigantescas.
- Invertimos mucho esfuerzo en medir información redundante para luego desecharla.
- Paradigma de la adquisición compresiva :
  1. Muestrear proporcionalmente a la **complejidad**.

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.
- En aplicaciones modernas como tratamiento de imágenes o vídeo, las dimensiones de las señales son gigantescas.
- Invertimos mucho esfuerzo en medir información redundante para luego desecharla.
- Paradigma de la adquisición compresiva :
  1. Muestrear proporcionalmente a la **complejidad**.
- Las muestras son una versión comprimida de la señal, medimos y comprimimos **a la vez**.

Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.
- En aplicaciones modernas como tratamiento de imágenes o vídeo, las dimensiones de las señales son gigantescas.
- Invertimos mucho esfuerzo en medir información redundante para luego desecharla.
- Paradigma de la adquisición compresiva :
  1. Muestrear proporcionalmente a la **complejidad**.
- Las muestras son una versión comprimida de la señal, medimos y comprimimos **a la vez**.
- Precio a pagar :

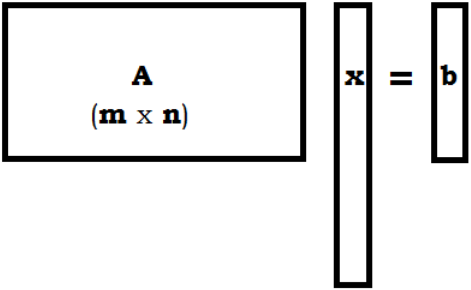
Motivación para la Adquisición Compresiva

- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.
- En aplicaciones modernas como tratamiento de imágenes o vídeo, las dimensiones de las señales son gigantescas.
- Invertimos mucho esfuerzo en medir información redundante para luego desecharla.
- Paradigma de la adquisición compresiva :
  1. Muestrear proporcionalmente a la **complejidad**.
- Las muestras son una versión comprimida de la señal, medimos y comprimimos **a la vez**.
- Precio a pagar :
  - Pequeño incremento en la tasa de compresión.

Motivación para la Adquisición Compresiva

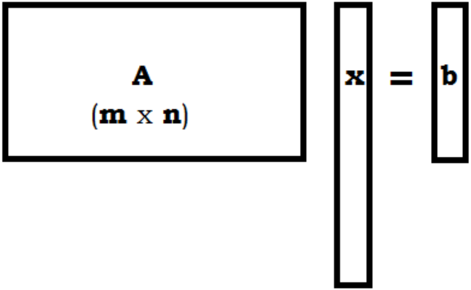
- Paradigma clásico de tratamiento de señales :
  1. Muestrear proporcionalmente al **ancho de banda**.
  2. Comprimir.
- En aplicaciones modernas como tratamiento de imágenes o vídeo, las dimensiones de las señales son gigantescas.
- Invertimos mucho esfuerzo en medir información redundante para luego desecharla.
- Paradigma de la adquisición compresiva :
  1. Muestrear proporcionalmente a la **complejidad**.
- Las muestras son una versión comprimida de la señal, medimos y comprimimos **a la vez**.
- Precio a pagar :
  - Pequeño incremento en la tasa de compresión.
  - Condiciones sobre el proceso de muestreo.

Sistema Indeterminado de Ecuaciones



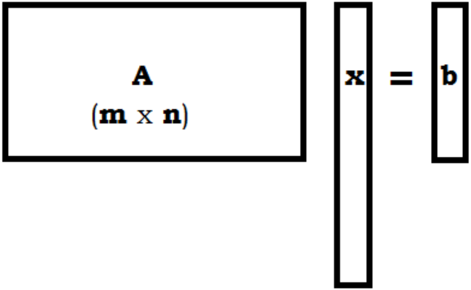
- La matriz **A** modela el submuestreo lineal de una señal **x**.

Sistema Indeterminado de Ecuaciones



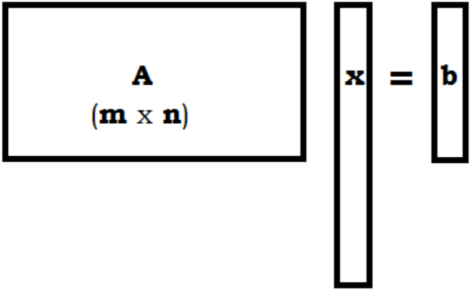
- La matriz **A** modela el submuestreo lineal de una señal **x**.
- Modelo útil para muchas aplicaciones.

Sistema Indeterminado de Ecuaciones



- La matriz **A** modela el submuestreo lineal de una señal **x**.
- Modelo útil para muchas aplicaciones.
- **Problema :** ¿Cuál es el **mínimo** número de muestras **m** necesario para reconstruir **x** ?

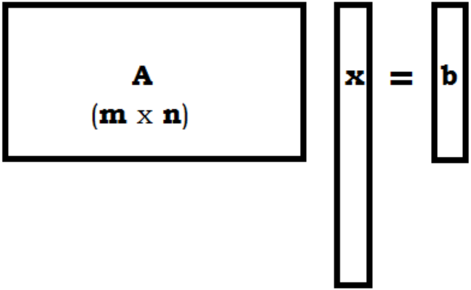
Sistema Indeterminado de Ecuaciones



- La matriz **A** modela el submuestreo lineal de una señal **x**.
- Modelo útil para muchas aplicaciones.
- **Problema :** ¿Cuál es el **mínimo** número de muestras **m** necesario para reconstruir **x** ?

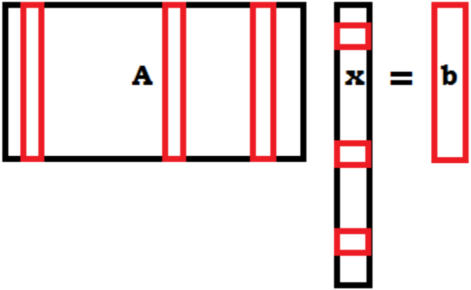
  m = **n** proporcional a la resolución (si A es invertible).

Sistema Indeterminado de Ecuaciones



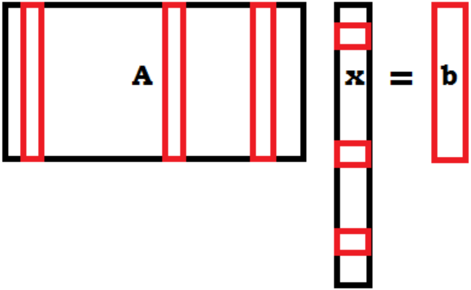
- **Hipótesis** : **x** es **s-disperso**, sólo un subconjunto **T** de **s** índices de **x** es distinto de cero.

Sistema Indeterminado de Ecuaciones



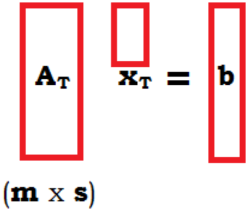
- **Hipótesis** : **x** es **s-disperso**, sólo un subconjunto **T** de **s** índices de **x** es distinto de cero.

Sistema Indeterminado de Ecuaciones



- **Hipótesis** : **x** es **s-disperso**, sólo un subconjunto **T** de **s** índices de **x** es distinto de cero.
- ¿Cuál es el **m** mínimo en este caso si conocemos **T** ?

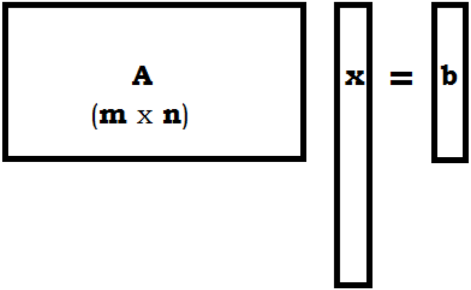
Sistema Indeterminado de Ecuaciones



- **Hipótesis** : **x** es **s-disperso**, sólo un subconjunto **T** de **s** índices de **x** es distinto de cero.
- ¿Cuál es el **m** mínimo en este caso si conocemos **T** ?

  m = **s** proporcional a la complejidad (si A_T es invertible).

Sistema Indeterminado de Ecuaciones



- **Hipótesis** : **x** es **s-disperso**, sólo un subconjunto **T** de **s** índices de **x** es distinto de cero.
- ¿Cuál es el **m** mínimo en este caso si conocemos **T**?
  m = **s** proporcional a la complejidad (si A_T es invertible).
- En muchas aplicaciones, la hipótesis es válida, pero **no conocemos T**.

Proceso de Adquisición Aleatorio

- Nuestro objetivo es reducir la dimensionalidad de la señal de manera **no adaptativa**.

Proceso de Adquisición Aleatorio

- Nuestro objetivo es reducir la dimensionalidad de la señal de manera **no adaptativa**.
- **Problema** : ¿Qué matrices $\mathbf{A}$ permiten hacer esto ?

Proceso de Adquisición Aleatorio

- Nuestro objetivo es reducir la dimensionalidad de la señal de manera **no adaptativa**.
- **Problema** : ¿Qué matrices $\mathbf{A}$ permiten hacer esto ?
- **Problema relacionado** : Reducir la dimensionalidad de varias señales conservando las distancias entre ellas.

Proceso de Adquisición Aleatorio

- Nuestro objetivo es reducir la dimensionalidad de la señal de manera **no adaptativa**.
- **Problema** : ¿Qué matrices **A** permiten hacer esto ?
- **Problema relacionado** : Reducir la dimensionalidad de varias señales conservando las distancias entre ellas.

  Según el **Lema de Johnson-Lindenstrauss** basta con proyectar de **manera aleatoria**.

Proceso de Adquisición Aleatorio

- Nuestro objetivo es reducir la dimensionalidad de la señal de manera **no adaptativa**.
- **Problema** : ¿Qué matrices **A** permiten hacer esto ?
- **Problema relacionado** : Reducir la dimensionalidad de varias señales conservando las distancias entre ellas.

  Según el **Lema de Johnson-Lindenstrauss** basta con proyectar de **manera aleatoria**.
- Utilizamos matrices aleatorias, por ejemplo $A_{ij} \sim \mathcal{N}(0,1)$.

Proceso de Adquisición Aleatorio

- Nuestro objetivo es reducir la dimensionalidad de la señal de manera **no adaptativa**.
- **Problema** : ¿Qué matrices **A** permiten hacer esto ?
- **Problema relacionado** : Reducir la dimensionalidad de varias señales conservando las distancias entre ellas.

  Según el **Lema de Johnson-Lindenstrauss** basta con proyectar de **manera aleatoria**.
- Utilizamos matrices aleatorias, por ejemplo $A_{ij} \sim \mathcal{N}(0,1)$.
- Sorprendentemente, esto nos permite reconstruir cualquier $\mathbf{x}$ $\mathbf{s}$-disperso con alta probabilidad a partir de $\mathbf{m} = C\ \mathbf{s} \log(\mathbf{n})$ muestras.

Reconstrucción Mediante Optimización

- A partir de $\mathbf{A}$ y $\mathbf{b}$, ¿cómo reconstruimos $\mathbf{x}$ sin conocer $\mathbf{T}$?

Reconstrucción Mediante Optimización

- A partir de **A** y **b**, ¿cómo reconstruimos **x** sin conocer **T**?
- Idealmente, resolveríamos el siguiente problema de optimización :

$$\min_{x} ||x||_0 \text{ sujeto a } Ax = b$$

$$||x||_0 = \text{número de elementos de } \mathbf{x} \text{ distintos de cero.}$$

Reconstrucción Mediante Optimización

- A partir de **A** y **b**, ¿cómo reconstruimos **x** sin conocer **T**?
- Idealmente, resolveríamos el siguiente problema de optimización :

$$\min_{x} ||x||_0 \text{ sujeto a } Ax = b$$

$$||x||_0 = \text{número de elementos de } \mathbf{x} \text{ distintos de cero.}$$

- Este problema es **NP-complejo** (*NP-hard*).

Reconstrucción Mediante Optimización

- A partir de **A** y **b**, ¿cómo reconstruimos **x** sin conocer **T** ?
- Idealmente, resolveríamos el siguiente problema de optimización :

$$\min_x ||x||_0 \text{ sujeto a } Ax = b$$

$||x||_0$ = número de elementos de **x** distintos de cero.

- Este problema es **NP-complejo** (*NP-hard*).
- En su lugar, resolvemos :

$$\min_x ||x||_1 \text{ sujeto a } Ax = b$$

$||x||_1 = \sum_{i=1}^{n} |x_i|$

Reconstrucción Mediante Optimización

- A partir de **A** y **b**, ¿cómo reconstruimos **x** sin conocer **T** ?
- Idealmente, resolveríamos el siguiente problema de optimización :

$$\min_x ||x||_0 \text{ sujeto a } Ax = b$$

$||x||_0$ = número de elementos de **x** distintos de cero.

- Este problema es **NP-complejo** (*NP-hard*).
- En su lugar, resolvemos :

$$\min_x ||x||_1 \text{ sujeto a } Ax = b$$

$||x||_1 = \sum_{i=1}^{n} |x_i|$

- Este problema es resoluble de manera **eficiente** (reformulable como un programa lineal).

Intuición Geométrica

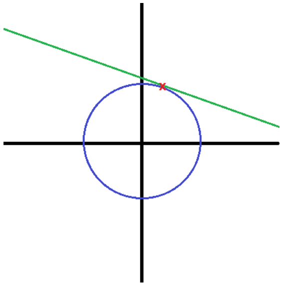
- La restricción $Ax = b$ define un hiperplano aleatorio de dimensión **m**.

Intuición Geométrica

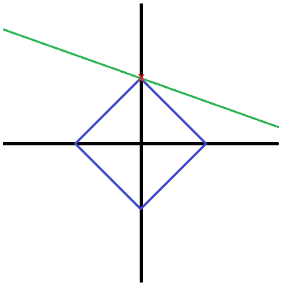
- La restricción $Ax = b$ define un hiperplano aleatorio de dimensión **m**.
- La solución es el punto en el que el hiperplano es tangente a la bola inducida por la norma que minimizamos.

Intuición Geométrica

- La restricción $Ax = b$ define un hiperplano aleatorio de dimensión **m**.
- La solución es el punto en el que el hiperplano es tangente a la bola inducida por la norma que minimizamos.
- Si la bola es puntiaguda, la solución estará próxima a los ejes.



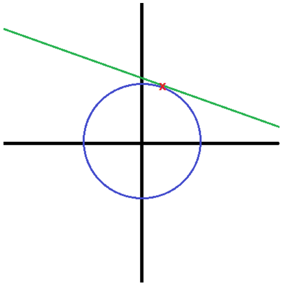
Norma ℓ_2 : $||x||_2 = \sqrt{\sum_{i=1}^{n} |x_i|}$



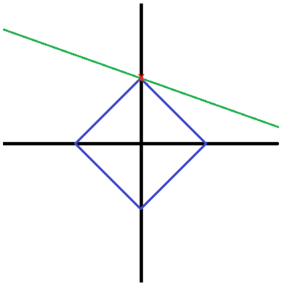
Norma ℓ_1 : $||x||_1 = \sum_{i=1}^{n} |x_i|$

Intuición Geométrica

- La restricción $Ax = b$ define un hiperplano aleatorio de dimensión **m**.
- La solución es el punto en el que el hiperplano es tangente a la bola inducida por la norma que minimizamos.
- Si la bola es puntiaguda, la solución estará próxima a los ejes.



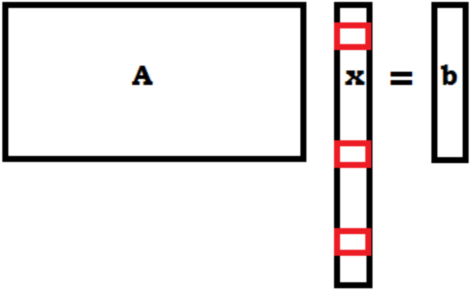
Norma $\ell_2 : ||x||_2 = \sqrt{\sum_{i=1}^{n} |x_i|}$



Norma $\ell_1 : ||x||_1 = \sum_{i=1}^{n} |x_i|$

- La bola ℓ_1 es la bola **convexa** más puntiaguda.

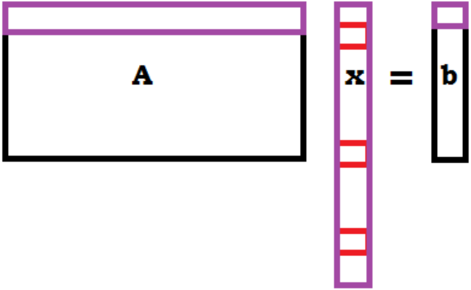
Muestras Aleatorizadas



- Interpretamos cada muestra como el producto de la señal dispersa **x** con un vector de muestreo ***a*** :

$$b_k = \langle a(k), x \rangle$$

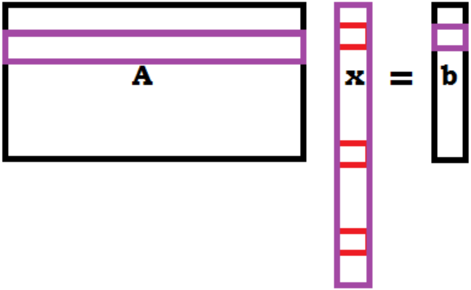
Muestras Aleatorizadas



- Interpretamos cada muestra como el producto de la señal dispersa **x** con un vector de muestreo ***a*** :

$$b_k = \langle a(k), x \rangle$$

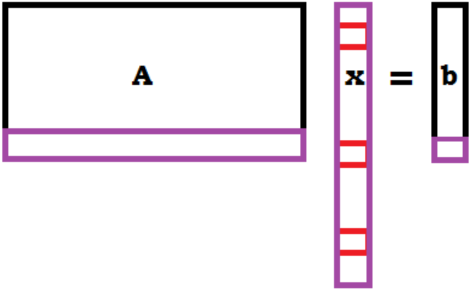
Muestras Aleatorizadas



- Interpretamos cada muestra como el producto de la señal dispersa **x** con un vector de muestreo ***a*** :

$$b_k = \langle a(k), x \rangle$$

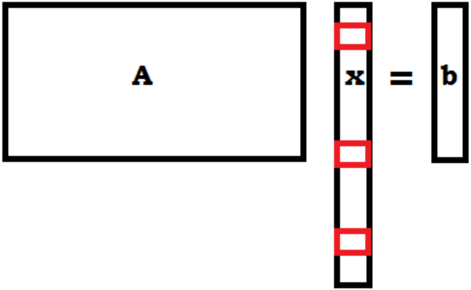
Muestras Aleatorizadas



- Interpretamos cada muestra como el producto de la señal dispersa $\mathbf{x}$ con un vector de muestreo $\boldsymbol{a}$:

$$b_k = \langle a(k), x \rangle$$

Muestras Aleatorizadas



- Interpretamos cada muestra como el producto de la señal dispersa **x** con un vector de muestreo ***a*** :

$$b_k = \langle a(k), x \rangle$$

- Reconstruímos minimizando la norma ℓ_1 dadas las muestras.

Reconstrucción por Optimización Convexa

- Utilizamos vectores aleatorios $\mathbf{a}$.

Reconstrucción por Optimización Convexa

- Utilizamos vectores aleatorios **a**.
- Deben cumplir dos condiciones :

$$\textbf{Incoherencia}: \quad \max_{1\le k\le n} |a_k|^2 \le \mu$$

$$\textbf{Isotropía}: \quad \mathbb{E}\left(aa^*\right) = \mathrm{I}$$

Reconstrucción por Optimización Convexa

- Utilizamos vectores aleatorios **a**.
- Deben cumplir dos condiciones :

Incoherencia : $\max_{1 \le k \le n} |a_k|^2 \le \mu$

Isotropía : $\mathbb{E}(aa^*) = \mathrm{I}$

Teorema [Candès, Plan 2010]

Un vector s-disperso x fijo es recuperable por minimización de la norma ℓ_1 con probabilidad $1 - 1/n - e^{-\beta}$ si el número de muestras satisface :

$$m \ge C\,\mu\,s\,\log n$$

Reconstrucción por Optimización Convexa

- Utilizamos vectores aleatorios **a**.
- Deben cumplir dos condiciones :

Incoherencia : $\max_{1 \leq k \leq n} |a_k|^2 \leq \mu$

Isotropía : $\mathbb{E}\left(aa^*\right) = \mathrm{I}$

Teorema [Candès, Plan 2010]

Un vector s-disperso x fijo es recuperable por minimización de la norma ℓ_1 con probabilidad $1 - 1/n - e^{-\beta}$ si el número de muestras satisface :

$$m \geq C\,\mu\,s\,\log n$$

- Muestreo proporcional a la complejidad si la incoherencia $\boldsymbol{\mu}$ es $\mathbf{O(1)}$ con alta probabilidad.

Incoherencia

- Si conociésemos $\mathbf{T}$, podríamos elegir cada vector de muestreo de tal manera que $b_k = x_k$ para todo $\mathbf{k}$ en $\mathbf{T}$.

Incoherencia

- Si conociésemos **T**, podríamos elegir cada vector de muestreo de tal manera que $\boldsymbol{b_k} = \boldsymbol{x_k}$ para todo **k** en **T**.
- Este muestreo es **altamente coherente**, la energía estaría concentrada en un solo coeficiente (incoherencia **O(n)**).

Incoherencia

- Si conociésemos $\mathbf{T}$, podríamos elegir cada vector de muestreo de tal manera que $\boldsymbol{b_k} = \boldsymbol{x_k}$ para todo $\mathbf{k}$ en $\mathbf{T}$.
- Este muestreo es **altamente coherente**, la energía estaría concentrada en un solo coeficiente (incoherencia $\mathbf{O(n)}$).
- Si **no** conocemos $\mathbf{T}$ esto es **desastroso**, con alta probabilidad necesitamos casi $\mathbf{n}$ muestras para recuperar la señal.

Incoherencia

- Si conociésemos **T**, podríamos elegir cada vector de muestreo de tal manera que $b_k = x_k$ para todo **k** en **T**.
- Este muestreo es **altamente coherente**, la energía estaría concentrada en un solo coeficiente (incoherencia **O(n)**).
- Si **no** conocemos **T** esto es **desastroso**, con alta probabilidad necesitamos casi **n** muestras para recuperar la señal.
- La energía de los vectores de muestreo debe estar **bien distribuida** para detectar coeficientes no nulos de **x** en **cualquier** posición.

Isotropía

- La incoherencia **no** es suficiente.

Isotropía

- La incoherencia **no** es suficiente.
- Por ejemplo, si **a** es constante, nunca podremos recontruir **x**.

Isotropía

- La incoherencia **no** es suficiente.
- Por ejemplo, si **a** es constante, nunca podremos recontruir **x**.
- Los vectores de muestreo deben cubrir **todas** las direcciones posibles.

Isotropía

- La incoherencia **no** es suficiente.
- Por ejemplo, si **a** es constante, nunca podremos recontruir **x**.
- Los vectores de muestreo deben cubrir **todas** las direcciones posibles.
- La condición de isotropía asegura que no favorecemos **ninguna dirección específica** al muestrear.

Vectores de Muestreo Válidos

- Vectores iid gaussianos $\mathcal{N}(0,\mathtt{I})$.

Vectores de Muestreo Válidos

- Vectores iid gaussianos $\mathcal{N}(0,\mathtt{I})$.
- Vectores con entradas de signo aleatorio ± 1 con prob. $1/2$.

Vectores de Muestreo Válidos

- Vectores iid gaussianos $\mathcal{N}(0,\mathtt{I})$.
- Vectores con entradas de signo aleatorio ± 1 con prob. $1/2$.
- Filas aleatorias de la Transformada Discreta de Fourier.

Vectores de Muestreo Válidos

- Vectores iid gaussianos $\mathcal{N}(0,\mathtt{I})$.
- Vectores con entradas de signo aleatorio ± 1 con prob. $1/2$.
- Filas aleatorias de la Transformada Discreta de Fourier.
- Filas aleatorias de cualquier transformada ortogonal incoherente.

Condición de Optimalidad

- Problema de optimización con restricciones lineales :

$$\min_x f(x) \text{ sujeto a } Ax = b$$

Condición de Optimalidad

- Problema de optimización con restricciones lineales :

$$\min_x f(x) \text{ sujeto a } Ax = b$$

- Si la función de coste es **diferenciable** y **convexa**, **x** es la solución si y sólo si el gradiente en **x** es **perpendicular** al núcleo de **A**.

Condición de Optimalidad

- Problema de optimización con restricciones lineales :

$$\min_{x} f(x) \text{ sujeto a } Ax = b$$

- Si la función de coste es **diferenciable** y **convexa**, **x** es la solución si y sólo si el gradiente en **x** es **perpendicular** al núcleo de **A**.
- La norma ℓ_1 no es diferenciable, pero la condición de optimalidad se puede generalizar.

Condición de Optimalidad

- Problema de optimización con restricciones lineales :

$$\min_x f(x) \text{ sujeto a } Ax = b$$

- Si la función de coste es **diferenciable** y **convexa**, **x** es la solución si y sólo si el gradiente en **x** es **perpendicular** al núcleo de **A**.
- La norma ℓ_1 no es diferenciable, pero la condición de optimalidad se puede generalizar.
- Un **subgradiente g** de una función **f** en **x** es un vector que satisface :

$$f(x+h) \geq f(x) + \langle g, h \rangle$$

Condición de Optimalidad

- Problema de optimización con restricciones lineales :

$$\min_x f(x) \text{ sujeto a } Ax = b$$

- Si la función de coste es **diferenciable** y **convexa**, **x** es la solución si y sólo si el gradiente en **x** es **perpendicular** al núcleo de **A**.
- La norma ℓ_1 no es diferenciable, pero la condición de optimalidad se puede generalizar.
- Un **subgradiente g** de una función **f** en **x** es un vector que satisface :

$$f(x+h) \geq f(x) + \langle g, h \rangle$$

- Si la función de coste es **convexa**, **x** es la solución si y sólo si existe un subgradiente en **x** **perpendicular** al núcleo de **A**.

Certificado Dual

- Basta con encontrar un subgradiente de la norma ℓ_1 en el **vector s-disperso original x** perpendicular al núcleo de **A**.

Certificado Dual

- Basta con encontrar un subgradiente de la norma ℓ_1 en el **vector s-disperso original x** perpendicular al núcleo de **A**.
- Llamamos a dicho vector un **certificado dual** porque está relacionado con el problema dual y certifica que tenemos una solución.

Certificado Dual

- Basta con encontrar un subgradiente de la norma ℓ_1 en el **vector s-disperso original x** perpendicular al núcleo de **A**.
- Llamamos a dicho vector un **certificado dual** porque está relacionado con el problema dual y certifica que tenemos una solución.
- Originalmente, esta era la parte más complicada de la demostración.

Certificado Dual

- Basta con encontrar un subgradiente de la norma ℓ_1 en el **vector s-disperso original x** perpendicular al núcleo de **A**.
- Llamamos a dicho vector un **certificado dual** porque está relacionado con el problema dual y certifica que tenemos una solución.
- Originalmente, esta era la parte más complicada de la demostración.
- Se puede simplificar enormemente aplicando una generalización de la Desigualdad de Chernoff para matrices.

Extensión a Matrices de Bajo Rango

- Estas ideas se pueden extender a la reconstrucción de matrices de bajo rango a partir de un pequeño número de coeficientes.

Extensión a Matrices de Bajo Rango

- Estas ideas se pueden extender a la reconstrucción de matrices de bajo rango a partir de un pequeño número de coeficientes.
- La reconstrucción sólo es posible si la energía de la matriz está muy distribuida (condición de **incoherencia**).

Extensión a Matrices de Bajo Rango

- Estas ideas se pueden extender a la reconstrucción de matrices de bajo rango a partir de un pequeño número de coeficientes.
- La reconstrucción sólo es posible si la energía de la matriz está muy distribuida (condición de **incoherencia**).
- Los coeficientes son medidos uniformemente al azar (condición de **isotropía**).

Extensión a Matrices de Bajo Rango

- Estas ideas se pueden extender a la reconstrucción de matrices de bajo rango a partir de un pequeño número de coeficientes.
- La reconstrucción sólo es posible si la energía de la matriz está muy distribuida (condición de **incoherencia**).
- Los coeficientes son medidos uniformemente al azar (condición de **isotropía**).
- Una matriz de bajo rango tiene **valores singulares dispersos**, por lo que minimizamos su norma ℓ_1.

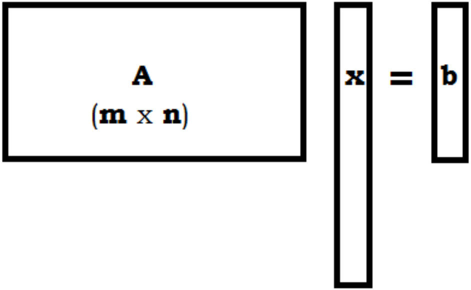
Extensión a Matrices de Bajo Rango

- Estas ideas se pueden extender a la reconstrucción de matrices de bajo rango a partir de un pequeño número de coeficientes.
- La reconstrucción sólo es posible si la energía de la matriz está muy distribuida (condición de **incoherencia**).
- Los coeficientes son medidos uniformemente al azar (condición de **isotropía**).
- Una matriz de bajo rango tiene **valores singulares dispersos**, por lo que minimizamos su norma ℓ_1.
- Esto es una norma matricial válida llamada **norma nuclear**.

Extensión a Matrices de Bajo Rango

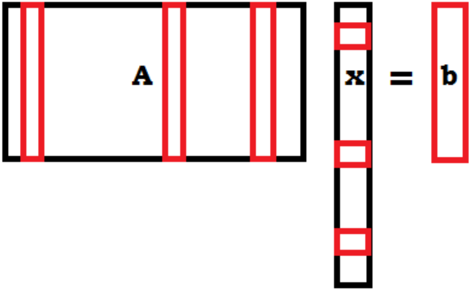
- Estas ideas se pueden extender a la reconstrucción de matrices de bajo rango a partir de un pequeño número de coeficientes.
- La reconstrucción sólo es posible si la energía de la matriz está muy distribuida (condición de **incoherencia**).
- Los coeficientes son medidos uniformemente al azar (condición de **isotropía**).
- Una matriz de bajo rango tiene **valores singulares dispersos**, por lo que minimizamos su norma ℓ_1.
- Esto es una norma matricial válida llamada **norma nuclear**.
- La demostración también se basa en la construcción de un **certificado dual** válido correspondiente a la matriz original.

Condicionamiento de Submatrices



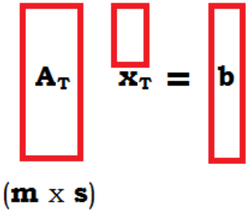
- En la práctica el proceso de adquisición añade **ruido** a las medidas **b**.

Condicionamiento de Submatrices



- En la práctica el proceso de adquisición añade **ruido** a las medidas **b**.
- Para una matriz **A** concreta, ¿que condición garantiza que podamos reconstruir cualquier **x** s-disperso de manera **estable**?

Condicionamiento de Submatrices



- En la práctica el proceso de adquisición añade **ruido** a las medidas **b**.
- Para una matriz **A** concreta, ¿que condición garantiza que podamos reconstruir cualquier **x** s-disperso de manera **estable**?
- Como mínimo, la matriz A_T debe estar bien condicionada para cualquier **T**.

Constante de Isometría Restringida

- **Constante de Isometría Restringida** : Mínima constante $\boldsymbol{\delta_s}$ tal que para **todo** vector **x** **s**-disperso :

$$(1-\delta_s)\,||x||_2^2 \leq ||Ax||_2^2 \leq (1+\delta_s)\,||x||_2^2$$

Constante de Isometría Restringida

- **Constante de Isometría Restringida** : Mínima constante $\boldsymbol{\delta_s}$ tal que para **todo** vector **x** **s**-disperso :

$$(1-\delta_s)\,||x||_2^2 \leq ||Ax||_2^2 \leq (1+\delta_s)\,||x||_2^2$$

Lema

Si $\boldsymbol{\delta_{2s}} < 1$, **A** es invertible cuando actúa sobre vectores **s**-dispersos.

Constante de Isometría Restringida

- **Constante de Isometría Restringida** : Mínima constante $\boldsymbol{\delta_s}$ tal que para **todo** vector **x** **s**-disperso :

$$(1-\delta_s)\,||x||_2^2 \leq ||Ax||_2^2 \leq (1+\delta_s)\,||x||_2^2$$

Lema

Si $\boldsymbol{\delta_{2s}} < 1$, **A** es invertible cuando actúa sobre vectores **s**-dispersos.

Demostración : Sea **y** tal que Ax=Ay :

$$||x-y||_2 \leq \frac{1}{\sqrt{1-\delta_{2s}}}\,||A\,(x-y)||_2 = 0$$

Constante de Isometría Restringida

- **Constante de Isometría Restringida** : Mínima constante $\boldsymbol{\delta_s}$ tal que para **todo** vector **x** **s**-disperso :

$$(1-\delta_s)\,||x||_2^2 \leq ||Ax||_2^2 \leq (1+\delta_s)\,||x||_2^2$$

Lema

Si $\boldsymbol{\delta_{2s}} < 1$, **A** es invertible cuando actúa sobre vectores **s**-dispersos.

Demostración : Sea **y** tal que Ax=Ay :

$$||x-y||_2 \leq \frac{1}{\sqrt{1-\delta_{2s}}}\,||A\,(x-y)||_2 = 0$$

- **No** implica que **A** se pueda invertir de manera **eficiente**.

Minimización de la Norma ℓ_1

Teorema [Candès, Romberg, Tao 2006]

Si $\boldsymbol{\delta_{2s}} < \sqrt{2} - 1$, todo vector s-disperso x es la solución del problema de optimización :

$$\min_{y} ||y||_1 \ \text{ sujeto a } Ay = Ax$$

Minimización de la Norma ℓ_1

Teorema [Candès, Romberg, Tao 2006]

Si $\delta_{2s} < \sqrt{2} - 1$, todo vector s-disperso x es la solución del problema de optimización :

$$\min_{y} ||y||_1 \text{ sujeto a } Ay = Ax$$

Acotar la Constante de Isometría Restringida de una matriz **garantiza** que es apta para la adquisición compresiva de **cualquier** señal dispersa.

Señales Compresibles y Medidas Ruidosas

- En la práctica :

Señales Compresibles y Medidas Ruidosas

- En la práctica :
  1. las señales rara vez son **exactamente** dispersas.

Señales Compresibles y Medidas Ruidosas

- En la práctica :
  1. las señales rara vez son **exactamente** dispersas.
  2. el proceso de adquisición añade **ruido** a las medidas.

Señales Compresibles y Medidas Ruidosas

- En la práctica :
  1. las señales rara vez son **exactamente** dispersas.
  2. el proceso de adquisición añade **ruido** a las medidas.
- En este caso resolvemos :

$$\min_{y} ||y||_1 \text{ sujeto a } ||Ay - b||_2 \leq \sigma$$

Señales Compresibles y Medidas Ruidosas

- En la práctica :
  1. las señales rara vez son **exactamente** dispersas.
  2. el proceso de adquisición añade **ruido** a las medidas.
- En este caso resolvemos :

$$\min_{y} ||y||_1 \text{ sujeto a } ||Ay - b||_2 \leq \sigma$$

- Definimos $\mathbf{x_s}$ como la mejor aproximación **s**-dispersa de **x**.

Señales Compresibles y Medidas Ruidosas

- En la práctica :
  1. las señales rara vez son **exactamente** dispersas.
  2. el proceso de adquisición añade **ruido** a las medidas.
- En este caso resolvemos :

$$\min_{y} ||y||_1 \text{ sujeto a } ||Ay - b||_2 \leq \sigma$$

- Definimos $\boldsymbol{x_s}$ como la mejor aproximación **s**-dispersa de **x**.
- Idealmente, el error de estimación al recuperar **x** sería proporcional al error de aproximación $||x - x_s||$ y al nivel de ruido $\boldsymbol{\sigma}$.

Señales Compresibles y Medidas Ruidosas

Teorema [Candès, Romberg, Tao 2006]

Si $\boldsymbol{\delta_{2s}} < \sqrt{2} - 1$, para cualquier señal **x** y ruido **z** tal que $||z||_2 \leq \sigma$, la solución del problema de optimización :

$$\hat{x} = \min_{y} ||y||_1 \ \text{ sujeto a } \ ||Ay - (Ax + z)||_2 \leq \sigma$$

satisface

$$||x - \hat{x}||_2 \leq C_0 \frac{||x - x_s||_1}{\sqrt{s}} + C_1 \sigma$$

Constante de Isometría Restringida de Matrices Aleatorias

- **Buenas noticias** : La RIC nos asegura la reconstrucción estable y eficiente de señales aproximadamente compresibles.

Constante de Isometría Restringida de Matrices Aleatorias

- **Buenas noticias** : La RIC nos asegura la reconstrucción estable y eficiente de señales aproximadamente compresibles.
- **Malas noticias** : Calcular la RIC de una matriz $\mathbf{A}$ concreta es NP-complejo.

Constante de Isometría Restringida de Matrices Aleatorias

- **Buenas noticias** : La RIC nos asegura la reconstrucción estable y eficiente de señales aproximadamente compresibles.
- **Malas noticias** : Calcular la RIC de una matriz **A** concreta es NP-complejo.
- Sin embargo, $\boldsymbol{\delta_{2s}} < \sqrt{2} - 1$ con alta probabilidad si :

Constante de Isometría Restringida de Matrices Aleatorias

- **Buenas noticias** : La RIC nos asegura la reconstrucción estable y eficiente de señales aproximadamente compresibles.
- **Malas noticias** : Calcular la RIC de una matriz **A** concreta es NP-complejo.
- Sin embargo, $\delta_{2s} < \sqrt{2} - 1$ con alta probabilidad si :
  - **A** es **gaussiana** ($A_{ij} \sim \mathcal{N}(0,1)$) o **binaria** ($A_{ij} = \pm 1$ con prob. 1/2) y $\boldsymbol{m = O\left(s \log\left(n/s\right)\right)}$

Constante de Isometría Restringida de Matrices Aleatorias

- **Buenas noticias** : La RIC nos asegura la reconstrucción estable y eficiente de señales aproximadamente compresibles.
- **Malas noticias** : Calcular la RIC de una matriz **A** concreta es NP-complejo.
- Sin embargo, $\delta_{2s} < \sqrt{2} - 1$ con alta probabilidad si :
  - **A** es **gaussiana** ($A_{ij} \sim \mathcal{N}(0,1)$) o **binaria** ($A_{ij} = \pm 1$ con prob. 1/2) y $\boldsymbol{m = O\left(s \log\left(n/s\right)\right)}$
  - **A** consta de filas **aleatorias** de la transformada discreta de Fourier y $\boldsymbol{m = O\left(s\left(\log n\right)^4\right)}$

Constante de Isometría Restringida de Matrices Aleatorias

- **Buenas noticias** : La RIC nos asegura la reconstrucción estable y eficiente de señales aproximadamente compresibles.
- **Malas noticias** : Calcular la RIC de una matriz **A** concreta es NP-complejo.
- Sin embargo, $\delta_{2s} < \sqrt{2} - 1$ con alta probabilidad si :
  - **A** es **gaussiana** ($A_{ij} \sim \mathcal{N}(0{,}1)$) o **binaria** ($A_{ij} = \pm 1$ con prob. 1/2) y $\boldsymbol{m = O\left(s \log\left(n/s\right)\right)}$
  - **A** consta de filas **aleatorias** de la transformada discreta de Fourier y $\boldsymbol{m = O\left(s \left(\log n\right)^4\right)}$
- Las matrices aleatorias nos permiten submuestrear proporcionalmente a la complejidad de la señal.

RIC de una Matriz Gaussiana

- Definición de valor singular mínimo de una matriz :

$$\sigma_{\text{min}} = \min_{||x||_2 \leq 1} ||Mx||_2$$

RIC de una Matriz Gaussiana

- Definición de valor singular mínimo de una matriz :

$$\sigma_{\min} = \min_{||x||_2 \leq 1} ||Mx||_2$$

- Debemos acotar el valor singular mínimo de **toda** submatriz de A con **s** columnas.

RIC de una Matriz Gaussiana

- Definición de valor singular mínimo de una matriz :

$$\sigma_{\min} = \min_{||x||_2 \leq 1} ||Mx||_2$$

- Debemos acotar el valor singular mínimo de **toda** submatriz de A con **s** columnas.
- Para ello :

RIC de una Matriz Gaussiana

- Definición de valor singular mínimo de una matriz :

$$\sigma_{\min} = \min_{||x||_2 \leq 1} ||Mx||_2$$

- Debemos acotar el valor singular mínimo de **toda** submatriz de A con **s** columnas.
- Para ello :
  1. Acotamos la probabilidad de que $\boldsymbol{\sigma_{\min}}$ supere $\boldsymbol{\sqrt{1-\delta_s}}$ para una submatriz fija.

RIC de una Matriz Gaussiana

- Definición de valor singular mínimo de una matriz :

$$\sigma_{\min} = \min_{||x||_2 \leq 1} ||Mx||_2$$

- Debemos acotar el valor singular mínimo de **toda** submatriz de A con **s** columnas.
- Para ello :
  1. Acotamos la probabilidad de que $\boldsymbol{\sigma_{\min}}$ supere $\boldsymbol{\sqrt{1-\delta_s}}$ para una submatriz fija.
  2. Aplicamos la desigualdad de Boole multiplicando por el número de submatrices con **s** columnas : $\boldsymbol{\binom{n}{s}}$.

Aplicaciones

- Geofísica (adquisición de datos sísmicos).

Aplicaciones

- Geofísica (adquisición de datos sísmicos).
- Imagen médica (resonancia magnética, tomografía por rayos X).

Aplicaciones

- Geofísica (adquisición de datos sísmicos).
- Imagen médica (resonancia magnética, tomografía por rayos X).
- Óptica (cámara de un solo píxel).

Aplicaciones

- Geofísica (adquisición de datos sísmicos).
- Imagen médica (resonancia magnética, tomografía por rayos X).
- Óptica (cámara de un solo píxel).
- Genética (diseño de microarrays de ADN).

Referencias

- Precedentes :

Referencias

- Precedentes :
  - Linear inversion of band-limited reflection seismograms
    Santosa, Symes 1986

Referencias

- Precedentes :
  - Linear inversion of band-limited reflection seismograms
    Santosa, Symes 1986
  - Atomic Decomposition by Basis Pursuit
    Chen, Donoho, Saunders 1995

Referencias

- Precedentes :
  - Linear inversion of band-limited reflection seismograms
    Santosa, Symes 1986
  - Atomic Decomposition by Basis Pursuit
    Chen, Donoho, Saunders 1995
  - Regression shrinkage and selection via the lasso
    Tibshirani 1996

Referencias

- Precedentes :
  - Linear inversion of band-limited reflection seismograms
    Santosa, Symes 1986
  - Atomic Decomposition by Basis Pursuit
    Chen, Donoho, Saunders 1995
  - Regression shrinkage and selection via the lasso
    Tibshirani 1996
- Demostración basada en certificado dual :

Referencias

- Precedentes :
  - Linear inversion of band-limited reflection seismograms
    Santosa, Symes 1986
  - Atomic Decomposition by Basis Pursuit
    Chen, Donoho, Saunders 1995
  - Regression shrinkage and selection via the lasso
    Tibshirani 1996
- Demostración basada en certificado dual :
  - Robust Uncertainty Principles : Exact Signal Reconstruction from Highly Incomplete. Frequency Information
    Candès, Romberg, Tao 2005

Referencias

- Precedentes :
  - Linear inversion of band-limited reflection seismograms
    Santosa, Symes 1986
  - Atomic Decomposition by Basis Pursuit
    Chen, Donoho, Saunders 1995
  - Regression shrinkage and selection via the lasso
    Tibshirani 1996
- Demostración basada en certificado dual :
  - Robust Uncertainty Principles : Exact Signal Reconstruction from Highly Incomplete. Frequency Information
    Candès, Romberg, Tao 2005
  - A probabilistic and RIPless theory of compressed sensing
    Candès, Plan 2010

Referencias

- Matrices de bajo rango :

Referencias

- Matrices de bajo rango :
  - Exact matrix completion via convex optimization
    Candès, Recht 2008

Referencias

- Matrices de bajo rango :
  - Exact matrix completion via convex optimization
    Candès, Recht 2008
  - Recovering low-rank matrices from few coefficients in any basis
    Gross 2011

Referencias

- Matrices de bajo rango :
  - Exact matrix completion via convex optimization
    Candès, Recht 2008
  - Recovering low-rank matrices from few coefficients in any basis
    Gross 2011
- Demostración basada en constante de isometría restringida :

Referencias

- Matrices de bajo rango :
  - Exact matrix completion via convex optimization
    Candès, Recht 2008
  - Recovering low-rank matrices from few coefficients in any basis
    Gross 2011
- Demostración basada en constante de isometría restringida :
  - Decoding by linear programming
    Candès, Tao 2006

Referencias

- Matrices de bajo rango :
  - Exact matrix completion via convex optimization
    Candès, Recht 2008
  - Recovering low-rank matrices from few coefficients in any basis
    Gross 2011
- Demostración basada en constante de isometría restringida :
  - Decoding by linear programming
    Candès, Tao 2006
- Otros enfoques :

Referencias

- Matrices de bajo rango :
  - Exact matrix completion via convex optimization
    Candès, Recht 2008
  - Recovering low-rank matrices from few coefficients in any basis
    Gross 2011
- Demostración basada en constante de isometría restringida :
  - Decoding by linear programming
    Candès, Tao 2006
- Otros enfoques :
  - Compressed Sensing
    Donoho 2005

Referencias

- Matrices de bajo rango :
  - Exact matrix completion via convex optimization
    Candès, Recht 2008
  - Recovering low-rank matrices from few coefficients in any basis
    Gross 2011
- Demostración basada en constante de isometría restringida :
  - Decoding by linear programming
    Candès, Tao 2006
- Otros enfoques :
  - Compressed Sensing
    Donoho 2005
  - Compressed Sensing over ℓ_p-balls : Minimax Mean Square Error
    Donoho, Johnstone, Maleki, Montanari 2010

Perspectivas Futuras

- Construcción determinística de isometrías restringidas.

Perspectivas Futuras

- Construcción determinística de isometrías restringidas.
- Relajación de las condiciones sobre el proceso de adquisición.

Perspectivas Futuras

- Construcción determinística de isometrías restringidas.
- Relajación de las condiciones sobre el proceso de adquisición.
- Extensión a señales dispersas en diccionarios redundantes.

Perspectivas Futuras

- Construcción determinística de isometrías restringidas.
- Relajación de las condiciones sobre el proceso de adquisición.
- Extensión a señales dispersas en diccionarios redundantes.
- Extensión a otros objetos de baja complejidad como tensores.

Gracias