Random projections

Optimization-Based Data Analysis
http://www.cims.nyu.edu/~cfgranda/pages/OBDA_spring16

Carlos Fernandez-Granda

2/29/2016
Introduction

Random projections in data analysis and signal processing

They preserve information embedded in low-dimensional subspaces of high-dimensional spaces

Non-adaptive compression, agnostic to specific data
Dimensionality reduction
  Principal component analysis
  Random projections

Compressed sensing
  Motivation: Magnetic resonance imaging
  Exact recovery
  Robustness

Sampling
  Nyquist-Shannon sampling theorem
  Compressive sampling
Dimensionality reduction

Projection of data onto lower-dimensional space

- Decreases computational cost of processing the data
- Allows to visualize (2D, 3D)

We will focus on linear projections
Linear projection

The linear projection of $x \in \mathbb{R}^n$ onto a subspace $S \subseteq \mathbb{R}^n$ of dimension $m \leq n$ is the solution to

$$\begin{align*}
\text{minimize} & \quad \|x - u\|_2 \\
\text{subject to} & \quad u \in S
\end{align*}$$

If the columns of $U$: $U_1, \ldots, U_m$ are an orthonormal basis of $S$

$$\mathcal{P}_S(x) = \sum_{i=1}^{m} \langle x, U_i \rangle U_i = UU^T x$$

To reduce the dimension we represent the signal using the coefficients

$$c := U^T x \in \mathbb{R}^m$$
Dimensionality reduction
   Principal component analysis
   Random projections

Compressed sensing
   Motivation: Magnetic resonance imaging
   Exact recovery
   Robustness

Sampling
   Nyquist-Shannon sampling theorem
   Compressive sampling
Adaptive projection

Data: $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_k$

Preprocessing: Centering the data

\[ x_i = \tilde{x}_i - \frac{1}{k} \sum_{i=1}^{k} \tilde{x}_i \]

Aim: Find directions of maximum variation
Principal component analysis (PCA)

1. Group the centered data in a data matrix $X$

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_k \end{bmatrix}$$

2. Compute the SVD of $X = U\Sigma V^T$

3. Extract the first $m$ left singular vectors

$$\hat{U} = \begin{bmatrix} U_1 & \cdots & U_m \end{bmatrix}$$

For any $n$-dimensional subspace $S'$

$$\sum_{i=1}^{k} \left\| P_{S'} x_i \right\|_2^2 \leq \sum_{i=1}^{k} \left\| \hat{U} \hat{U}^T x_i \right\|_2^2$$
Example: 2D data

\[ \frac{\sigma_1}{\sqrt{k}} = 0.705 \quad \frac{\sigma_2}{\sqrt{k}} = 0.690 \]
Example: 2D data

\[ \frac{\sigma_1}{\sqrt{k}} = 0.9832 \quad \frac{\sigma_2}{\sqrt{k}} = 0.3559 \]
Example: 2D data

\[
\frac{\sigma_1}{\sqrt{k}} = 1.3490 \quad \frac{\sigma_2}{\sqrt{k}} = 0.1438
\]

\[u_1\]
\[u_2\]
Example

Seeds from three different varieties of wheat: Kama, Rosa and Canadian

Dimensions:

- Area
- Perimeter
- Compactness
- Length of kernel
- Width of kernel
- Asymmetry coefficient
- Length of kernel groove
Projection onto two first PCs
Projection onto two last PCs
Dimensionality reduction
   Principal component analysis
   Random projections

Compressed sensing
   Motivation: Magnetic resonance imaging
   Exact recovery
   Robustness

Sampling
   Nyquist-Shannon sampling theorem
   Compressive sampling
Non-adaptive projections

PCA requires processing all of the data before projecting

Idea: Project onto random $m$-dimensional subspace

Not optimal, but more computationally efficient

Approximate projection: Multiply by a random matrix $A \in \mathbb{R}^{m \times n}$
Approximate projection onto two random directions
**Johnson-Lindenstrauss lemma**

Random function $f$ preserves distance between points

$$f(x) := \frac{1}{\sqrt{m}}Ax$$

$A$ is an $m \times n$ matrix with iid Gaussian entries with mean 0 and variance 1 (can be generalized to Bernouilli $\pm 1$ entries)

Fix $x_1, \ldots, x_k \in \mathbb{R}^n$. For any $x_i \neq x_j$

$$(1 - \epsilon) \|x_i - x_j\|_2^2 \leq \|f(x_i) - f(x_j)\|_2^2 \leq (1 + \epsilon) \|x_i - x_j\|_2^2$$

with probability at least $\frac{1}{k}$ as long as

$$m \geq \frac{8 \log (k)}{\epsilon^2}$$
Result for fixed vector

For any fixed vector $v \in \mathbb{R}^n$

$$(1 - \epsilon) \|v\|_2^2 \leq \frac{1}{m} \|Av\|_2^2 \leq (1 + \epsilon) \|v\|_2^2$$

with probability at least

$$1 - 2 \exp \left( -\frac{m\epsilon^2}{8} \right)$$

Combining this with the union bound yields the result
Proof of result for fixed vector

Apply concentration bound on chi-square random variable $Z$ with $m$ degrees of freedom

$$Z := \sum_{i=1}^{m} X_i^2$$

$X_1, \ldots, X_m$ are Gaussian with mean 0 and variance 1 and independent

For any $\epsilon > 0$ we have

$$P (Z > m (1 + \epsilon)) \leq \exp \left( - \frac{m\epsilon^2}{8} \right)$$

$$P (Z < m (1 - \epsilon)) \leq \exp \left( - \frac{m\epsilon^2}{2} \right)$$
Dimensionality reduction
  Principal component analysis
  Random projections

Compressed sensing
  Motivation: Magnetic resonance imaging
  Exact recovery
  Robustness

Sampling
  Nyquist-Shannon sampling theorem
  Compressive sampling
Dimensionality reduction
   Principal component analysis
   Random projections

Compressed sensing
   Motivation: Magnetic resonance imaging
   Exact recovery
   Robustness

Sampling
   Nyquist-Shannon sampling theorem
   Compressive sampling
Magnetic resonance imaging
Magnetic resonance imaging

**Data:** Samples from spectrum

**Problem:** Sampling is time consuming (annoying, kids move . . . )

Images are **compressible** (sparse in wavelet basis)

Can we recover compressible signals from less data?
Idea

By now (hopefully) we know that \( \ell_1 \)-norm induces sparsity

1. Undersample data
2. Solve the optimization problem

\[
\begin{align*}
\text{minimize} & \quad \| \text{wavelet transform of estimate} \|_1 \\
\text{subject to} & \quad \text{frequency samples of estimate} = \text{data}
\end{align*}
\]
Regular vs random undersampling

Minimum $\ell_2$-norm estimate
Minimum $\ell_1$-norm estimate
Dimensionality reduction
   Principal component analysis
   Random projections

Compressed sensing
   Motivation: Magnetic resonance imaging
   Exact recovery
   Robustness

Sampling
   Nyquist-Shannon sampling theorem
   Compressive sampling
Random samples

1. Undersample the spectrum randomly

Signal

Spectrum

---

Data
\( \ell_1 \)-norm minimization

2. Solve the optimization problem

\[
\begin{align*}
\text{minimize} & \quad ||\text{estimate}||_1 \\
\text{subject to} & \quad \text{frequency samples of estimate} = \text{data}
\end{align*}
\]
Underdetermined system of equations
Underdetermined system of equations
Ax = y where $A \in \mathbb{R}^{m \times n}$ and $m < n$, infinite solutions!
Exact recovery

Assumption: There exists a signal $x \in \mathbb{R}^n$ with $s$ nonzeros such that

$$Ax = y$$

for a random $A \in \mathbb{R}^{m \times n}$ (random Fourier, Gaussian iid, Bernoulli $\pm 1$, ...)

Exact recovery: If the number of measurements satisfies

$$m \geq C's \log n$$

the solution of the problem

$$\text{minimize} \quad \|\tilde{x}\|_1 \quad \text{subject to} \quad A\tilde{x} = y$$

is the original signal with probability at least $1 - \frac{1}{n}$.
Incoherent measurements

**Generalization:** Random rows $U_j$ from orthonormal basis $U$

**Coherence:**

$$\mu(U) := \sqrt{n} \max_{1 \leq i \leq n, 1 \leq j \leq m} |U_j e_i|$$

Exact recovery is achieved with high probability if

$$m \geq C' \mu(U) s \log n$$

**Random Fourier:** $\mu(F) = 1$
The dual problem is equal to

maximize \( y^T \tilde{v} \)
subject to \( \| A^T \tilde{v} \|_\infty \leq 1 \)
A dual certificate $v \in \mathbb{R}^m$ associated to $x$ is equal to

\[
\left(A^T v\right)_i = \text{sign}(x_i) \quad \text{if } x_i \neq 0
\]

\[
\|\left(A^T v\right)_i\|_{\infty} < 1 \quad \text{if } x_i = 0
\]

Feasible for dual problem, corresponding cost-function value equals

\[
y^T v = \|x\|_1
\]

By weak duality $x$ must be a solution
By the definition of $v$

$$q := A^T v$$

is a subgradient of the $\ell_1$ norm at $x$ and for any $h$ such that $Ah = 0$

$$h^T q = 0$$

This also implies that $x$ is a solution

If $A_T$ (where $T$ is the support of $x$) is injective, $x$ is the unique solution
Proof of exact recovery

Prove that dual certificate exists for any $s$-sparse $x$

Idea: Choose vector that interpolates the sign and has minimum $\ell_2$ norm

\[
\begin{align*}
\text{minimize} & \quad \|\tilde{v}\|_2 \\
\text{subject to} & \quad A_T^T \tilde{v} = \text{sign} (x_T)
\end{align*}
\]

Closed-form solution $v_{\ell_2} = A_T \left( A_T^T A_T \right)^{-1} \text{sign} (x_T)$

We need to prove that $q_{\ell_2} := A^T v_{\ell_2}$ satisfies

\[
\begin{align*}
(q_{\ell_2})_T &= \text{sign} (x_T) \\
\| (q_{\ell_2})_{T^c} \|_\infty &< 1
\end{align*}
\]
Random Fourier measurements

Tough stuff, we will prove the result for Gaussian measurements
Bounds on singular values of Gaussian submatrix

Fix a support $T$, $|T| \leq s$

For any unit-norm vector $x$ with support $T$

$$1 - \epsilon \leq \frac{1}{\sqrt{m}} \|Ax\|_2^2 \leq 1 + \epsilon$$

with probability at least

$$1 - 2 \left( \frac{12}{\epsilon} \right)^s \exp \left( -\frac{m\epsilon^2}{32} \right)$$
Bounds on singular values of Gaussian submatrix

Setting $\epsilon = 1/2$ gives

$$1 - \frac{1}{2} \leq \frac{1}{\sqrt{m}} \|Ax\|_2 \leq 1 + \frac{1}{2}$$

with probability at least

$$1 - \exp \left( -\frac{Cm}{s} \right)$$

for some constant $C$
Bound on dual certificate

Minimum singular value of $A_T$

$$\sigma_{\text{min}} (A_T) \geq \frac{\sqrt{m}}{2}$$

with probability $1 - \exp \left( - \frac{Cm}{s} \right)$

This implies $A_T^T A_T$ is invertible so

$$(q_{\ell_2})_T = A_T^T A_T \left( A_T^T A_T \right)^{-1} \text{sign} (x_T) = \text{sign} (x_T)$$
Bound on dual certificate

To bound \((q_{\ell_2})_{T^c}\), for each \(i \in T^c\) we define

\[
(q_{\ell_2})_i = A_i^T A_T \left( A_T^T A_T \right)^{-1} \text{sign} \left( x_T \right)
= A_i^T w
\]

\(A_i\) and \(w\) are independent

By the bound on \(\sigma_{\min}(A_T)\)

\[
\|w\|_2 \leq \frac{\|\text{sign} \left( x_T \right)\|_2}{\sigma_{\min}(A_T)} \leq 2 \sqrt{\frac{s}{m}}
\]

with probability \(1 - \exp \left(-\frac{Cm}{s}\right)\)
Bound on dual certificate

Conditioned on $w$, $A_i^T w$ is Gaussian with mean 0 and variance $||w||_2^2$

$$P \left( \left| A_i^T w \right| \geq 1 \mid w = w' \right) \leq P \left( |u| > \frac{1}{||w'||_2} \right)$$

$$\leq 2 \exp \left( -\frac{1}{2 ||w'||_2^2} \right)$$

Where $u$ has mean 0 and variance 1

For $\mathcal{E} := \left\{ ||w||_2 \leq 2 \sqrt{\frac{s}{m}} \right\}$ this implies

$$P \left( \left| A_i^T w \right| \geq 1 \mid \mathcal{E} \right) \leq 2 \exp \left( -\frac{m}{8s} \right)$$
Bound on dual certificate

Finally

\[
P \left( \left| A_i^T w \right| \geq 1 \right) \leq P \left( \left| A_i^T w \right| \geq 1 \mid \mathcal{E} \right) + P \left( \mathcal{E}^c \right) \\
\leq \exp \left( -\frac{Cm}{s} \right) + 2 \exp \left( -\frac{m}{8s} \right)
\]

If the number of measurements satisfies

\[
m \geq C' s \log n
\]

we have exact recovery with probability \(1 - \frac{1}{n}\) by the union bound
Let $\mathcal{X}_T$ be the set of unit-norm vectors $x$ with support $T$

**Aim:** Prove that for any $x \in \mathcal{X}_T$

$$(1 - \epsilon) \leq \frac{1}{\sqrt{m}} ||Ax||_2 \leq (1 + \epsilon)$$

With probability $1 - 2 \exp \left( - \frac{me^2}{8} \right)$ for any fixed unit-norm vector $v$

$$(1 - \epsilon) \leq \frac{1}{m} ||Av||_2^2 \leq (1 + \epsilon)$$

We apply this result on an $\epsilon$-net of $\mathcal{X}_T$
\( \mathcal{N}_\epsilon \subseteq \mathcal{X} \) is an \( \epsilon \)-net of \( \mathcal{X} \) if for every \( y \in \mathcal{X} \) there is \( x \in \mathcal{N}_\epsilon \) such that 
\[
\|x - y\|_2 \leq \epsilon.
\]

The covering number \( \mathcal{N}(\mathcal{X}, \epsilon) \) of a set \( \mathcal{X} \) at scale \( \epsilon \) is the minimal cardinality of an \( \epsilon \)-net of \( \mathcal{X} \).

The covering number of \( \mathcal{X}_T \) is
\[
\mathcal{N}(\mathcal{X}_T, \epsilon) \leq \left( \frac{3}{\epsilon} \right)^s
\]
Proof of bounds on singular values

By the union bound and the bound for fixed vectors

\[ \left| \frac{1}{m} \| Au \|_2^2 - 1 \right| > \frac{\epsilon}{2} \]

for some \( u \in \mathcal{N}(X, \epsilon/4) \) with probability at most

\[ 2 \left( \frac{12}{\epsilon} \right)^s \exp \left( -\frac{m\epsilon^2}{32} \right) \]
Proof of bounds on singular values

Assume that for all \( u \in \mathcal{N}(\mathcal{X}, \epsilon/4) \)

\[
1 - \frac{\epsilon}{2} \leq \frac{1}{\sqrt{m}} \|Au\|_2 \leq 1 + \frac{\epsilon}{2}
\]

Define \( \alpha \) as the smallest number such that for all \( x \in \mathcal{X}_T \)

\[
\frac{1}{\sqrt{m}} \|Ax\|_2 \leq 1 + \alpha
\]

For any \( x \in \mathcal{X}_T \), there is a \( u \in \mathcal{N}(\mathcal{X}, \epsilon/4) \) such that

\[
\frac{1}{\sqrt{m}} \|Ax\|_2 \leq \frac{1}{\sqrt{m}} \left( \|Au\|_2 + \|A(x - u)\|_2 \right)
\]

\[
\leq 1 + \frac{\epsilon}{2} + \frac{(1 + \alpha) \epsilon}{4}
\]
We conclude

$$\alpha \leq \frac{3\epsilon}{4 - \epsilon} \leq \epsilon$$

$$\frac{1}{\sqrt{m}} \|Ax\|_2 \geq \frac{1}{\sqrt{m}} (\|Au\|_2 - \|A(x - u)\|_2)$$

$$\geq 1 - \frac{\epsilon}{2} - \frac{(1 + \epsilon)\epsilon}{4}$$

$$\geq 1 - \epsilon$$
Dimensionality reduction
   Principal component analysis
   Random projections

Compressed sensing
   Motivation: Magnetic resonance imaging
   Exact recovery
   Robustness

Sampling
   Nyquist-Shannon sampling theorem
   Compressive sampling
Is the problem well posed?

\[ \text{Spectrum} \]
Is the problem well posed?

Spectrum
Is the problem well posed?
Is the problem well posed?

Aim: Study effect of measurement operator on sparse vectors
Is the problem well posed?

Equivalently, is the submatrix always well conditioned?
Is the problem well posed?

Equivalently, is the submatrix always well conditioned?
Restricted isometry property (RIP)

For any $s$-sparse vector $x$

$$(1 - \epsilon_s) \|x\|_2 \leq \frac{1}{\sqrt{m}} \|Ax\|_2 \leq (1 + \epsilon_s) \|x\|_2$$

with probability at least $1 - 2 \exp(-C_1 n)$ if

$$m \geq \frac{C_1 s}{\epsilon^2 s} \log \left( \frac{n}{s} \right)$$

2s-RIP implies that for any $s$-sparse signals $x_1, x_2$

$$\|y_2 - y_1\|_2 \geq (1 - \epsilon_{2s}) \|x_2 - x_1\|_2$$
Robustness

Noisy data

\[ y = Ax + z \quad \text{where } \|z\|_2 \leq \epsilon_0 \]

Relaxed problem

\[
\begin{align*}
\text{minimize} & \quad \|\tilde{x}\|_1 \\
\text{subject to} & \quad \|A\tilde{x} - y\|_2 \leq \epsilon_0
\end{align*}
\]
Robustness

If $x$ is $s$-sparse under the RIP, solution $\hat{x}$ satisfies

$$||\hat{x} - x||_2 \leq C_0 \epsilon_0$$

If $x$ is not sparse

$$||\hat{x} - x||_2 \leq C_0 \epsilon_0 + C_1 \frac{||x - x_s||_1}{\sqrt{s}}$$

$x_s$ contains $s$ entries of $x$ with largest magnitude
Proof of the RIP

For a fixed support $T$,

$$(1 - \epsilon) \|x\|_2 \leq \frac{1}{\sqrt{m}} \|Ax\|_2 \leq (1 + \epsilon) \|x\|_2$$

for any $x$ with support $T$ with probability at least

$$1 - 2 \left( \frac{12}{\epsilon} \right)^s \exp \left( -\frac{m\epsilon^2}{32} \right)$$
Proof of the RIP

Number of possible supports

$$\binom{n}{s} \leq \left( \frac{en}{s} \right)^s$$

By the union bound the result holds with probability at least

$$1 - 2 \left( \frac{en}{s} \right)^s \left( \frac{12}{\epsilon} \right)^s \exp \left( -\frac{m\epsilon^2}{32} \right)$$

$$= 1 - \exp \left( \log 2 + s + s \log \left( \frac{n}{s} \right) + s \log \left( \frac{12}{\epsilon} \right) - \frac{m\epsilon^2}{2} \right)$$

$$\leq 1 - \frac{C_2}{n}$$

as long as

$$m \geq \frac{C_1 s \log \left( \frac{n}{s} \right)}{\epsilon^2}$$
Dimensionality reduction
   Principal component analysis
   Random projections

Compressed sensing
   Motivation: Magnetic resonance imaging
   Exact recovery
   Robustness

Sampling
   Nyquist-Shannon sampling theorem
   Compressive sampling
Dimensionality reduction
   Principal component analysis
   Random projections

Compressed sensing
   Motivation: Magnetic resonance imaging
   Exact recovery
   Robustness

Sampling
   Nyquist-Shannon sampling theorem
   Compressive sampling
Sampling problem

Aim: Estimate bandlimited signal $g \in L_2([0, 1])$

$$g(t) := \sum_{k=-f}^{f} c_k \exp(i2\pi kt)$$

from samples

$$g(0), g\left(\frac{1}{n}\right), g\left(\frac{2}{n}\right), \ldots, g\left(\frac{n-1}{n}\right)$$

Questions:

1. At what rate do we need to sample?
2. How do we recover the signal from the samples?
Sampling problem

Signal

Spectrum
Sampling problem
Notation

\[ a_{-f:f} (t)_k := \begin{bmatrix}
\exp (-i2\pi (-f) t) \\
\exp (-i2\pi (-f + 1) t) \\
\vdots \\
\exp (-i2\pi (f - 1) t) \\
\exp (-i2\pi ft)
\end{bmatrix} \]

\[ g(t) := \sum_{k=-f}^{f} c_k \exp (i2\pi kt) = a_{-f:f} (t)^* c \]
Data

$n$ equations, $2f + 1$ unknowns

\[
F^* c = \begin{bmatrix}
    a_{-f:f}(0)^*
    \\ a_{-f:f}(1/n)^*
    \\ a_{-f:f}(2/n)^*
    \\ \vdots
    \\ a_{-f:f}(n-1/n)^*
\end{bmatrix}
\quad
\begin{bmatrix}
g(0)
    \\ g(1/n)
    \\ g(2/n)
    \\ \vdots
    \\ g(n-1/n)
\end{bmatrix}
\]

Sampling rate $n \geq 2f + 1$
Recovery

If \( n = 2f + 1 \), the vectors

\[
\frac{1}{\sqrt{n}} a_{-f:f} (0), \quad \frac{1}{\sqrt{n}} a_{-f:f} \left( \frac{1}{n} \right), \ldots, \quad \frac{1}{\sqrt{n}} a_{-f:f} \left( \frac{n-1}{n} \right)
\]

form an orthonormal basis, so

\[
c = \frac{1}{n} FF^* c = \frac{1}{n} F \begin{bmatrix} g(0) \\ g\left( \frac{1}{n} \right) \\ g\left( \frac{2}{n} \right) \\ \vdots \\ g\left( \frac{n-1}{n} \right) \end{bmatrix} = \frac{1}{n} \sum_{j=0}^{n} g\left( \frac{j}{n} \right) a_{-f:f} \left( \frac{j}{n} \right)
\]
Periodized sinc or Dirichlet kernel

\[ D(t) := \frac{1}{n} \sum_{k=-f}^{f} e^{-i 2\pi k t} = \frac{\sin(\pi nt)}{n \sin(\pi t)} \]
Interpolation with weighted sincs!

\[ g(t) = a_{-f:f}(t) \ast c \]

\[ = \frac{1}{n} \sum_{j=0}^{n} g\left(\frac{j}{n}\right) a_{-f:f}(t) \ast a_{-f:f}\left(\frac{j}{n}\right) \]

\[ = \sum_{j=0}^{n} g\left(\frac{j}{n}\right) D\left(t - \frac{j}{n}\right) \]

\[ D\left(t - \frac{j}{n}\right) = \frac{1}{n} \sum_{k=-f}^{f} e^{-i2\pi k\left(t - \frac{j}{n}\right)} \]

\[ = \frac{1}{n} a_{-f:f}(t) \ast a_{-f:f}\left(\frac{j}{n}\right) \]
Recovery
Nyquist-Shannon sampling theorem

**Condition:** Sampling rate \( \geq \) twice the highest frequency

**Recovery:** Interpolation with sinc kernel

Just linear algebra!
Dimensionality reduction
  Principal component analysis
  Random projections

Compressed sensing
  Motivation: Magnetic resonance imaging
  Exact recovery
  Robustness

Sampling
  Nyquist-Shannon sampling theorem
  Compressive sampling
Sparse spectrum

**Aim:** Estimate a signal $g \in L_2 ([0, 1])$ with a sparse spectrum

$$g (t) := \sum_{k \in S} c_k \exp (i2\pi kt)$$

- Maximum frequency: $f$
- Number of sinusoids: $s$

How many measurements do we need?

- Nyquist-Shannon: $2f + 1$
- Compressed sensing: $\mathcal{O} (s \log (2f + 1))$
Sparse spectrum
Linear estimation

$n$ equations, $2f + 1$ unknowns

\[ F^* c = \begin{bmatrix}
    a_{-f:f}(0)^* \\
    a_{-f:f}(\frac{1}{n})^* \\
    a_{-f:f}(\frac{2}{n})^* \\
    \vdots \\
    a_{-f:f}(\frac{n-1}{n})^*
\end{bmatrix} c = \begin{bmatrix}
    g(0) \\
    g\left(\frac{1}{n}\right) \\
    g\left(\frac{2}{n}\right) \\
    \vdots \\
    g\left(\frac{n-1}{n}\right)
\end{bmatrix} \]

Measurements: $2f + 1$

Recovery: Sinc interpolation
Compressive sampling

$m$ equations, $2f + 1$ unknowns

\[
F^* c = \begin{bmatrix}
    a_{f:f}(0)^* \\
    a_{f:f}(\frac{1}{n})^* \\
    a_{f:f}(\frac{2}{n})^* \\
    \vdots \\
    a_{f:f}(\frac{n-1}{n})^*
\end{bmatrix}
\begin{bmatrix}
    g(0) \\
    g(\frac{1}{n}) \\
    g(\frac{2}{n}) \\
    \vdots \\
    g(\frac{n-1}{n})
\end{bmatrix}
\]

Measurements: \( m \geq C s \log (2f + 1) \) (random undersampling)

Recovery: \( \ell_1 \)-norm minimization to compute \( c \)
Recovery

Linear estimate

Compressive sampling