情報基礎数理学特選
有限距離空間の離散幾何学とアルゴリズム
(5) 線形方程式系の疎解

岡本 吉央

北陸先端科学技術大学院大学

2011 年 1 月 12 日

"最終更新：2011/01/13 10:04"

この講義の主題

次の問題に対する考察

線形方程式系の疎解探索

入力
- 行列 $A \in \mathbb{R}^{m \times n}$， ベクトル $b \in \mathbb{R}^m$， 自然数 r

出力
- 方程式 $Ax = b$ の解 $x \in \mathbb{R}^n$ で $|\mathrm{supp}(x)| \leq r$ を満たすもの

定義：台

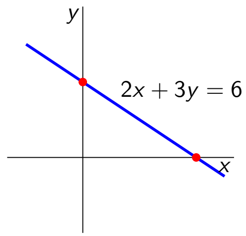
ベクトル $x \in \mathbb{R}^n$ の台 (support) とは

$$\mathrm{supp}(x) = \{i \in \{1, \ldots, n\} \mid x_i \neq 0\}$$

のこと

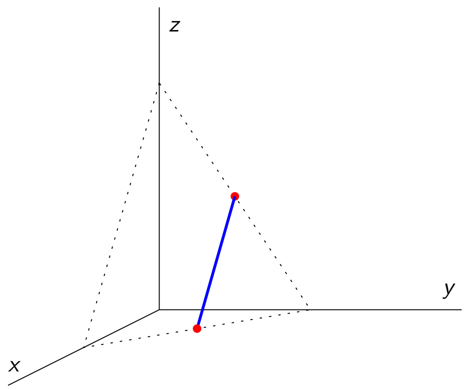
例 1

- $n = 2, m = 1, r = 1$
- 方程式 $2x + 3y = 6$ の解で $|\mathrm{supp}((x, y))| \leq 1$ のもの



例 2

- $n = 3, m = 2, r = 2$
- 方程式 $6x + 3y + 2z = 12, y = 2$ の解で $|\mathrm{supp}((x, y, z))| \leq 2$ のもの



どうしてこんな問題を考えるのか？

多くの応用があるから

(1) 回帰分析

(2) ノイズ除去

(3) 圧縮センシング

注：これらの応用では，A, r を自分で決められる

どうしてこんな問題を考えるのか？ (1)：回帰分析 (1)

- ブラックボックス $B \colon \mathbb{R}^n \to \mathbb{R}$ をモデル化して，推定したい

$$\begin{array}{ccccc} x_i & \longrightarrow & \boxed{B} & \longrightarrow & y_i \\ \cap & & & & \cap \\ \mathbb{R}^n & & (i = 1, \dots, m) & & \mathbb{R} \end{array}$$

- モデル化：$B(x) = w^\top x + \varepsilon$ (線形モデル)
    - $w \in \mathbb{R}^n$：回帰係数
    - $\varepsilon \in \mathbb{R}$：誤差
- 推定：誤差 2 乗和最小化による w の推定 (最小 2 乗法)

$$\inf \sum_{i=1}^{m} (y_i - w^\top x_i)^2$$

- 次を満たす w が存在すれば，それは最適解
    - すべての $i = 1, \ldots, m$ に対して $y_i = w^\top x_i$ (線形方程式系)

どうしてこんな問題を考えるのか？ (1)：回帰分析 (2)

- 「少ない変数による説明が望ましい」として最適化問題を変更 (r は適当に与える)

$$\inf \left\{ \sum_{i=1}^{m} (y_i - w^\top x_i)^2 \;\middle|\; |\mathrm{supp}(w)| \leq r \right\}$$

- つまり，次の 2 つの条件を満たす $w \in \mathbb{R}^n$ が欲しい
    - すべての $i = 1, \ldots, m$ に対して $y_i = w^\top x_i$ (線形方程式系)
    - $|\mathrm{supp}(w)| \leq r$ (疎性)

どうしてこんな問題を考えるのか？ (2)：ノイズ除去

- 観測データの通信

観測データ	$\boxed{m \in \mathbb{R}^k}$	
符号化	$\downarrow$	
送信データ	$\boxed{u = G^\top m \in \mathbb{R}^n}$	$(G^\top \in \mathbb{R}^{n \times k})$ 生成行列
通信	$\downarrow$	
受信データ	$\boxed{r = u + e \in \mathbb{R}^n}$	$(e \in \mathbb{R}^n)$ ノイズ

- ノイズに影響を受けた成分は少ないと仮定 $(|\mathrm{supp}(e)| \leq t)$
- 受信データから送信データを復元 (誤り訂正)
  - $HG^\top = O$ となる行列 $H \in \mathbb{R}^{\ell \times n}$ を考える　(検査行列)
  - $Hr = H(G^\top m + e) = HG^\top m + He = He$
- よって，次を満たす e が分かれば送信データ u を復元できる
  - $He = Hr$　(線形方程式系)
  - $|\mathrm{supp}(e)| \leq t$　(疎性)

どうしてこんな問題を考えるのか？ (3)：圧縮センシング

圧縮センシング (compressed sensing) の問題設定

- 疎な信号 $x \in \mathbb{R}^n$ の観測 $(|\mathrm{supp}(x)| \leq r)$
  - 普通の観測：n 個の成分を独立に観測
  - 「圧縮」観測：k 個のベクトル $a_1, \ldots, a_k \in \mathbb{R}^n$ を用意して $b_1 = a_1^\top x, \ldots, b_k = a_k^\top x$ を観測
- 問題：$b_1, \ldots, b_k$ から x が復元できるか？
- 復元した信号 $y \in \mathbb{R}^n$ が満たしていて欲しい条件
  - すべての $i = 1, \ldots, k$ に対して $a_i^\top y = b_i$　(線形方程式系)
  - $|\mathrm{supp}(y)| \leq r$　(疎性)

疎解探索問題の難しさ

線形方程式系の疎解探索

入力
- 行列 $A \in \mathbb{R}^{m \times n}$，ベクトル $b \in \mathbb{R}^m$，自然数 r

出力
- 方程式 $Ax = b$ の解 $x \in \mathbb{R}^n$ で $|\mathrm{supp}(x)| \leq r$ を満たすもの

事実　(Garey, Johnson '79)

線形方程式系の疎解探索問題は NP 困難

$\leadsto$ ヒューリスティックなアルゴリズムを考える

基底追跡

基底追跡 (basis pursuit, BP)

入力

- 行列 $A \in \mathbb{R}^{m \times n}$，ベクトル $b \in \mathbb{R}^m$

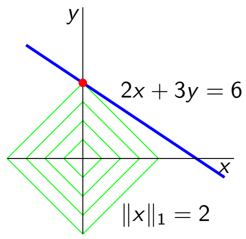
問題

- $\inf\{\|x\|_1 \mid Ax = b\}$

これは線形計画問題であり (演習問題)，多項式時間で解ける

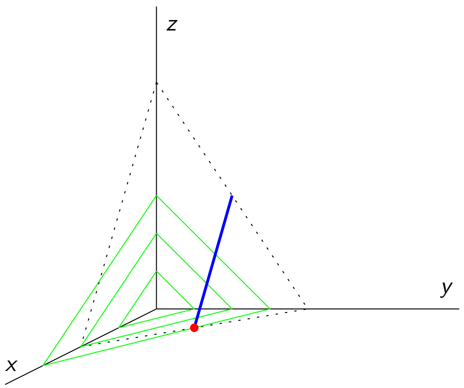
例 1

- $\inf\{|x| + |y| \mid 2x + 3y = 6\}$

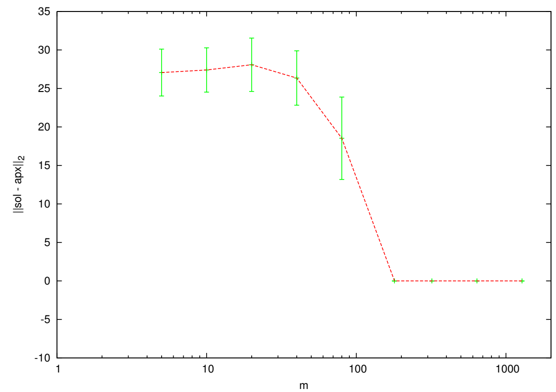


例 2

- $\inf\{|x| + |y| + |z| \mid 6x + 3y + 2z = 12, y = 2\}$



計算機実験



- 詳細は後掲
- $n = 2000, r = 20$：固定
- $m = 5, 10, 20, 40, 80, 160, 320, 640, 1280$：変化
- 疎解と基底追跡の最適解の差 (ユークリッド距離) の平均をプロット

線形方程式系の疎解と基底追跡

例と計算機実験からの観察

- 2 つの例では，基底追跡の最適解が疎解になっている
- 計算機実験では，m が小さくないとき，基底追跡の最適解が疎解になっている
- 計算機実験では，基底追跡の最適解が疎解ではないときもある

今から行うこと

- 基底追跡によって疎解が得られる場合がいつなのか，を考察する (理論的解析)

ここからの流れ

- 行列の制限付等長性 (RIP) を定義
- 次を証明 (詳細は後掲)
  - A が RIP を満たす，かつ，疎解が存在
    $\Rightarrow$ BP の最適解が疎解となる
  - A が JL 補題の射影行列，かつ，m が小さくない
    $\Rightarrow$ A が RIP を満たす

<u>注</u> (再)：先に見た応用では，A を自分で決められる

制限付等長性

m, n, t：自然数 ($m \leq n, t \leq n$)

定義：制限付等長性 (restricted isometry property, RIP)

行列 $A \in \mathbb{R}^{m \times n}$ が *t-制限付等長性* (t-restricted isometry property, t-RIP) を持つとは，
ある $\varepsilon \geq 0$ が存在して，$|\mathrm{supp}(x)| \leq t$ を満たす任意の $x \in \mathbb{R}^n$ に対して，

$$(1-\varepsilon)\|x\|_2 \leq \|Ax\|_2 \leq (1+\varepsilon)\|x\|_2$$

を満たすこと
このとき，ε を *t-制限付等長定数* (t-restricted isometry constant) と呼ぶ

解釈：疎なベクトルに対して，A は ℓ_2-ノルムをほとんど保存する

RIP ⇒ BP が疎解を与える

定理 5.1 (Candès, Tao '05, Rudelson, Vershynin '05, Donoho '06)

m, n, r：自然数，$A \in \mathbb{R}^{m \times n}$，$b \in \mathbb{R}^m$

A が $3r$-制限付等長性を持つ (定数は $\varepsilon > 0$)，かつ $Ax = b$ かつ $|\mathrm{supp}(x)| \leq r$ を満たすベクトル $x \in \mathbb{R}^n$ が唯一に存在する ⇒ その x は BP の唯一の最適解

定理 5.1 の証明 (1)

仮定は成り立つとする

- x が BP の唯一の最適解ではないと仮定して矛盾を導く
- y：BP の最適解 ($x \neq y$)
  - $Ax = b = Ay$
  - $\|x\|_1 \geq \|y\|_1$
- $e = y - x$ とする ($e \neq 0$)
  - 特に，$Ae = A(y-x) = Ay - Ax = 0$ であり，ゆえに，$\|Ae\|_2 = 0$
  - 最終的には，$e \neq 0$ に矛盾することを導く

定理 5.1 の証明 (2)

主張 1

$$\sum_{i \in \mathrm{supp}(x)} |e_i| \geq \sum_{i \notin \mathrm{supp}(x)} |e_i|$$

証明：$\|x\|_1 \geq \|y\|_1 = \|e + x\|_1$

$$\begin{aligned}
&= \sum_{i \in \mathrm{supp}(x)} |e_i + x_i| + \sum_{i \notin \mathrm{supp}(x)} |e_i| && (\because x_i = 0 \text{ for } i \notin \mathrm{supp}(x))\\
&\geq \sum_{i \in \mathrm{supp}(x)} (|x_i| - |e_i|) + \sum_{i \notin \mathrm{supp}(x)} |e_i| && (\because \text{三角不等式})\\
&= \sum_{i \in \mathrm{supp}(x)} |x_i| - \sum_{i \in \mathrm{supp}(x)} |e_i| + \sum_{i \notin \mathrm{supp}(x)} |e_i|\\
&= \|x\|_1 - \sum_{i \in \mathrm{supp}(x)} |e_i| + \sum_{i \notin \mathrm{supp}(x)} |e_i| && (\because x_i = 0 \text{ for } i \notin \mathrm{supp}(x))
\end{aligned}$$

□

定理 5.1 の証明 (3)

- $Z = \{1, \dots, n\} \setminus \mathrm{supp}(x)$ とする ($|Z| = n - |\mathrm{supp}(x)| \geq n - r$)
- Z を $k = \lceil (n-r)/2r \rceil$ 個のブロック $B_1, \dots, B_k$ に分割する
  - $|B_1| = \dots = |B_{k-1}| = 2r$，$|B_k| \leq 2r$
  - $i \in B_j, i' \in B_{j'}, j < j' \Rightarrow |e_i| \geq |e_{i'}|$

  つまり，$|e_i|$ が大きい方から順に $2r$ 個ずつ $B_1, B_2, \dots$ に入れていき，最終的な余りが B_k に入る

定理 5.1 の証明 (4)

主張 2

任意の $j \in \{2, \dots, k\}$ に対して， $\sqrt{\sum_{i \in B_j} |e_i|^2} \le \frac{1}{\sqrt{2r}} \sum_{i' \in B_{j-1}} |e_{i'}|$

証明：$j \in \{2, \dots, k\}$ を固定

- $i \in B_j, i' \in B_{j-1} \Rightarrow |e_i| \le |e_{i'}|$　($\because B_j$ の構成法)
- $|e_i| \le \frac{1}{|B_{j-1}|} \sum_{i' \in B_{j-1}} |e_{i'}| = \frac{1}{2r} \sum_{i' \in B_{j-1}} |e_{i'}|$
- $\sum_{i \in B_j} |e_i|^2 \le \sum_{i \in B_j} \left(\frac{1}{2r} \sum_{i' \in B_{j-1}} |e_{i'}| \right)^2 \le 2r \frac{1}{(2r)^2} \left(\sum_{i' \in B_{j-1}} |e_{i'}| \right)^2$　□

定理 5.1 の証明 (5)

主張 3

$$\sum_{j=2}^{k} \sqrt{\sum_{i \in B_j} |e_i|^2} \le \frac{1}{\sqrt{2}} \sqrt{\sum_{i' \in \mathrm{supp}(x)} |e_{i'}|^2}$$

証明：任意の $z \in \mathbb{R}^n$ に対して $\|z\|_1 \le \sqrt{n}\|z\|_2$ なので (演習問題)，

$$\begin{aligned}
\sum_{j=2}^{k} \sqrt{\sum_{i \in B_j} |e_i|^2} &\le \sum_{j=2}^{k} \frac{1}{\sqrt{2r}} \sum_{i' \in B_{j-1}} |e_{i'}| && \text{(主張 2)} \\
&\le \frac{1}{\sqrt{2r}} \sum_{j=1}^{k} \sum_{i' \in B_j} |e_{i'}| = \frac{1}{\sqrt{2r}} \sum_{i' \notin \mathrm{supp}(x)} |e_{i'}| \\
&\le \frac{1}{\sqrt{2r}} \sum_{i' \in \mathrm{supp}(x)} |e_{i'}| && \text{(主張 1)} \\
&\le \frac{1}{\sqrt{2}} \sqrt{\sum_{i' \in \mathrm{supp}(x)} |e_{i'}|^2} && \square
\end{aligned}$$

定理 5.1 の証明 (6)

- 添字集合 X に対して，ベクトル $e|_X \in \mathbb{R}^n$ を

$$(e|_X)_i = \begin{cases} e_i & (i \in X \text{ のとき}) \\ 0 & (\text{それ以外のとき}) \end{cases}$$

  と定義すると，
  - $e = \sum_{j=1}^{k} e|_{B_j} + e|_{\mathrm{supp}(x)} = e|_{B_1 \cup \mathrm{supp}(x)} + \sum_{j=2}^{k} e|_{B_j}$
  - $\|e|_{B_j}\|_2^2 = \sum_{i \in B_j} |e_i|^2$
- このとき，主張 3 を言い換えると

$$\sum_{j=2}^{k} \|e|_{B_j}\|_2 \le \frac{1}{\sqrt{2}} \|e|_{\mathrm{supp}(x)}\|_2$$

定理 5.1 の証明 (7)

$$\begin{aligned}
0 = \|Ae\|_2 &= \left\| A \left(e|_{B_1 \cup \mathrm{supp}(x)} + \sum_{j=2}^{k} e|_{B_j} \right) \right\|_2 \\
&\ge \|A(e|_{B_1 \cup \mathrm{supp}(x)})\|_2 - \sum_{j=2}^{k} \|A(e|_{B_j})\|_2 && \text{(三角不等式)} \\
&\ge (1-\varepsilon)\|e|_{B_1 \cup \mathrm{supp}(x)}\|_2 - \sum_{j=2}^{k} (1+\varepsilon)\|e|_{B_j}\|_2 && (A \text{ の } 3r\text{-制限付等長性}) \\
&\ge (1-\varepsilon)\|e|_{B_1 \cup \mathrm{supp}(x)}\|_2 - \frac{1+\varepsilon}{\sqrt{2}}\|e|_{\mathrm{supp}(x)}\|_2 && \text{(主張 3)} \\
&\ge (1-\varepsilon)\|e|_{\mathrm{supp}(x)}\|_2 - \frac{1+\varepsilon}{\sqrt{2}}\|e|_{\mathrm{supp}(x)}\|_2 \\
&= \left(1 - \varepsilon - \frac{1+\varepsilon}{\sqrt{2}} \right) \|e|_{\mathrm{supp}(x)}\|_2 \ge 0 && (\because \varepsilon \text{ が十分小さい})
\end{aligned}$$

よって，$\|e|_{\mathrm{supp}(x)}\|_2 = 0$ (すなわち，$e|_{\mathrm{supp}(x)} = 0$)

定理 5.1 の証明 (8)

$$0 = \|e|_{\mathrm{supp}(x)}\|_1 = \sum_{i \in \mathrm{supp}(x)} |e_i| \geq \sum_{i \notin \mathrm{supp}(x)} |e_i| \geq 0 \quad \text{(主張 1)}$$

よって，$e = 0$ となり，これは $e \neq 0$ に矛盾 □

ここからの流れ

- 行列の制限付等長性 (RIP) を定義 (済)
- 次を証明 (詳細は後掲)
  - A が RIP を満たす，かつ，疎解が存在 $\Rightarrow$ BP の最適解が疎解となる (済)
  - A が JL 補題の射影行列，かつ，m が小さくない $\Rightarrow$ A が RIP を満たす

注 (再)：先に見た応用では，A を自分で決められる

A が JL 補題の行列，かつ，m が小さくない $\Rightarrow$ A は RIP を満たす

m, n：自然数

定理 5.2 (Candès, Tao '05, Rudelson, Vershynin '05, Donoho '06)

設定：
- $A \in \mathbb{R}^{m \times n}$：各成分が独立に $\mathcal{N}(0,1)$ に従う
- $B = \frac{1}{\sqrt{m}} A$
- $\varepsilon > 0$：十分小さな定数
- ある定数 C に対して，$1 \leq r \leq n/C$ かつ $m \geq Cr \ln \frac{n}{r}$

結論：
- ε のみに依存する定数 $c > 0$ が存在して

$$\Pr[B \text{ が } 3r\text{-制限付等長性 (定数は } \varepsilon) \text{ を満たす}] \geq 1 - \exp(-cm)$$

つまり，JL 補題で使われた行列は高い確率で制限付等長性を満たす

JL 補題の証明で使った補題 (復習)

補題 4.1 の帰結

ε のみに依存する定数 $c_0 > 0$ が存在して，
任意の $x \in \mathbb{R}^n$ に対して

$$\Pr[(1-\varepsilon)\|x\|_2 \leq \|Bx\|_2 \leq (1+\varepsilon)\|x\|_2] \geq 1 - \exp(-c_0 m)$$

証明すべきこと

ε のみに依存する定数 $c > 0$ が存在して

$$\Pr\left[\begin{array}{l} |\mathrm{supp}(x)| \leq 3r \text{ を満たす任意の } x \text{ に対して} \\ (1-\varepsilon)\|x\|_2 \leq \|Bx\|_2 \leq (1+\varepsilon)\|x\|_2 \end{array}\right] \geq 1 - \exp(-cm)$$

- 「任意の x に対して」が確率の外にあるか中にあるか
- 疎性に関する条件がないかあるか

定理 5.2 の証明の流れ

証明の方針

- 「任意の x に対して」が確率の外にあるので，中に入れたい
- 「任意の x」を「ある有限集合 N の任意の要素 x」に変えたい
- これにより，合併上界 (union bound) が適用できるようになる

証明の流れ

1. 有限稠密集合 N を定義
2. 空間全体 $\mathbb{R}^n$ と有限稠密集合 N の関係を考察
3. 全体は，矛盾を導いて証明

有限稠密集合

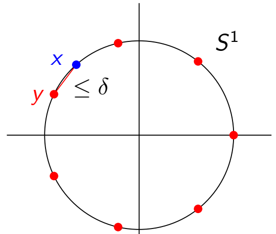
定義：単位球面

$$S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\} \subseteq \mathbb{R}^n$$

$\delta > 0$：固定

定義：稠密集合

集合 $N \subseteq S^{n-1}$ が δ-稠密 (δ-dense) であるとは，
$\forall\, x \in S^{n-1}, \exists\, y \in N$: $\|x - y\|_2 \leq \delta$



有限稠密集合と $(1+\varepsilon)$-埋め込み

補題 5.3

$\varepsilon > 0$：十分小さい，$N \subseteq S^{n-1}$：ε-稠密集合，$B \in \mathbb{R}^{m \times n}$

$$\forall\, y \in N\colon\ 1 - \varepsilon \leq \|By\|_2 \leq 1 + \varepsilon$$
$$\Downarrow$$
$$\forall\, x \in \mathbb{R}^n\colon\ (1 - 3\varepsilon)\|x\|_2 \leq \|Bx\|_2 \leq (1 + 3\varepsilon)\|x\|_2$$

証明：

- $x \in S^{n-1}$ に対して証明すれば十分
  - そうできたとする
  - $\tilde{x} \in \mathbb{R}^n$ に対して，$\tilde{x}/\|\tilde{x}\|_2 \in S^{n-1}$
  - $(1-3\varepsilon)\|\tilde{x}/\|\tilde{x}\|_2\|_2 \leq \|B(\tilde{x}/\|\tilde{x}\|_2)\|_2 \leq (1+3\varepsilon)\|\tilde{x}/\|\tilde{x}\|_2\|_2$
  - $(1-3\varepsilon)\|\tilde{x}\|_2 \leq \|B\tilde{x}\|_2 \leq (1+3\varepsilon)\|\tilde{x}\|_2$

補題 5.3 の証明 (続 1)

証明すること：任意の $x \in S^{n-1}$ に対して，$\|Bx\|_2 \leq 1 + 3\varepsilon$

- $M = \max\{\|Bx\|_2 \mid x \in S^{n-1}\}$ とする
- $x_0 \in S^{n-1}$：$M = \|Bx_0\|_2$ を満たすとする
- $y_0 \in N$：$\|x_0 - y_0\|_2 \leq \varepsilon$ を満たすとする
- 計算：$M = \|Bx_0\|_2 = \|B(x_0 - y_0 + y_0)\|_2 = \|B(x_0 - y_0) + By_0\|_2$

$$\begin{aligned}
&\leq \|B(x_0 - y_0)\|_2 + \|By_0\|_2 && \text{(三角不等式)}\\
&= \left\| \left(B \frac{x_0 - y_0}{\|x_0 - y_0\|_2} \right) \|x_0 - y_0\|_2 \right\|_2 + \|By_0\|_2\\
&\leq M\|x_0 - y_0\|_2 + 1 + \varepsilon\\
&\leq M\varepsilon + 1 + \varepsilon
\end{aligned}$$

- $\therefore M \leq (1+\varepsilon)/(1-\varepsilon) \leq 1 + 3\varepsilon$ $\qquad$ ($\because \varepsilon$ は十分小さい)
- $\therefore \|Bx\|_2 \leq M \leq 1 + 3\varepsilon$

補題 5.3 の証明 (続 2)

証明すること：任意の $x \in S^{n-1}$ に対して，$\|Bx\|_2 \geq 1 - 3\varepsilon$

- $y \in N$: $\|x - y\|_2 \leq \varepsilon$ を満たすとする
- 計算：$\|Bx\|_2 \geq \|By\|_2 - \|B(x-y)\|_2$ (三角不等式)

$$
\begin{aligned}
&= \|By\|_2 - \left\| B \frac{x-y}{\|x-y\|_2} \|x-y\|_2 \right\| \\
&\geq 1 - \varepsilon - (1+3\varepsilon)\varepsilon \\
&= 1 - 2\varepsilon - 3\varepsilon^2 \\
&\geq 1 - 3\varepsilon \qquad (\because \varepsilon \text{ は十分小さい})
\end{aligned}
$$

□

小さな稠密集合の存在性

補題 5.4

任意の $\varepsilon \in (0, 1]$ に対して，ε-稠密集合 $N \subseteq S^{n-1}$ で

$$|N| \leq \left(\frac{4}{\varepsilon}\right)^n$$

を満たすものが存在する

補題の証明で必要となる概念

定義：球

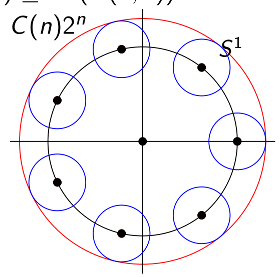
$(\mathbb{R}^n, \|\cdot\|_p)$ において，中心を $c \in \mathbb{R}^n$，半径を $r \geq 0$ とする球 (ball) とは，

$$B_p(c, r) = \{x \in \mathbb{R}^n \mid \|x - c\|_p \leq r\}.$$

事実：$\mathrm{vol}(B_2(c, r)) = C(n) r^n$
($C : \mathbb{R} \to \mathbb{R}$ はある関数で，$C(n) \leq O(1/\sqrt{n})^n$)

補題 5.4 の証明

- 次の性質を持つ集合 $N \subseteq S^{n-1}$ を考える
  - $\forall\, x \in S^{n-1} \setminus N$, $\exists\, y \in N$: $\|x - y\|_2 \leq \varepsilon$
  - $\forall\, y, y' \in N$: $\|y - y'\|_2 > \varepsilon$
- N は ε-稠密集合であり，そのような N は存在する
- $\forall\, y, y' \in N$: $B_2(y, \varepsilon/2) \cap B_2(y', \varepsilon/2) = \emptyset$
- $\forall\, y \in N$: $B_2(y, \varepsilon/2) \subseteq B_2(0, 1+\varepsilon/2) \subseteq B_2(0, 2)$
- $\therefore |N| \mathrm{vol}(B_2(0, \varepsilon/2)) \leq \mathrm{vol}(B(0, 2))$
- $\therefore |N| C(n)(\varepsilon/2)^n \leq C(n) 2^n$
- $\therefore |N| \leq (4/\varepsilon)^n$



□

定理 5.2 の証明の流れ (再掲)

証明の方針

- 「任意の x に対して」が確率の外にあるので，中に入れたい
- 「任意の x」を「ある有限集合 N の任意の要素 x」に変えたい
- これにより，合併上界 (union bound) が適用できるようになる

証明の流れ

1. 有限稠密集合 N を定義
2. 空間全体 $\mathbb{R}^n$ と有限稠密集合 N の関係を考察
3. 全体は，矛盾を導いて証明

定理 5.2 の証明 (1)

- $t = 3r$ とする（$t \le n$ と仮定してよい（C を十分大きな定数として））

$$
\Pr\left[\begin{array}{l}|\mathrm{supp}(x)| \le t \text{ を満たすある } x \in \mathbb{R}^n \text{ に対して} \\ (1-\varepsilon)\|x\|_2 \le \|Bx\|_2 \le (1+\varepsilon)\|x\|_2 \text{ が不成立}\end{array}\right]
$$
$$
\le \sum_{\substack{T\subseteq\{1,\dots,n\},\\ |T|\le t}} \Pr\left[\begin{array}{l}\mathrm{supp}(x) \subseteq T \text{ を満たすある } x \in \mathbb{R}^n \text{ に対して} \\ (1-\varepsilon)\|x\|_2 \le \|Bx\|_2 \le (1+\varepsilon)\|x\|_2 \text{ が不成立}\end{array}\right]
$$
$$
= \sum_{\substack{T\subseteq\{1,\dots,n\},\\ |T|\le t}} \Pr\left[\begin{array}{l}\text{ある } \tilde{x} \in \mathbb{R}^{|T|} \text{ に対して} \\ (1-\varepsilon)\|\tilde{x}\|_2 \le \|B_T\tilde{x}\|_2 \le (1+\varepsilon)\|\tilde{x}\|_2 \text{ が不成立}\end{array}\right]
$$

記法

$B_T \in \mathbb{R}^{m\times|T|}$：$T$ に対応する B の列からなる行列

定理 5.2 の証明 (2)

$T \subseteq \{1,\dots,n\}$, $|T| \le t$：固定

- $N_T \subseteq S^{|T|-1}$：ε-稠密集合（ただし，$|N_T| \le (4/\varepsilon)^{|T|}$）

補題 5.3 の帰結

$\exists\, \tilde{x} \in \mathbb{R}^{|T|}$: $(1-\varepsilon)\|\tilde{x}\|_2 \le \|B_T\tilde{x}\|_2 \le (1+\varepsilon)\|\tilde{x}\|_2$ が不成立 $\Rightarrow$
$\exists\, y \in N_T$: $(1-\varepsilon/3)\|y\|_2 \le \|B_T y\|_2 \le (1+\varepsilon/3)\|y\|_2$ が不成立

$$
\Pr\left[\begin{array}{l}\text{ある } \tilde{x} \in \mathbb{R}^{|T|} \text{ に対して} \\ (1-\varepsilon)\|\tilde{x}\|_2 \le \|B_T\tilde{x}\|_2 \le (1+\varepsilon)\|\tilde{x}\|_2 \text{ が不成立}\end{array}\right]
$$
$$
\le \Pr\left[\begin{array}{l}\text{ある } y \in N_T \text{ に対して} \\ (1-\varepsilon/3)\|y\|_2 \le \|B_T y\|_2 \le (1+\varepsilon/3)\|y\|_2 \text{ が不成立}\end{array}\right]
$$
$$
\le \sum_{y\in N_T} \Pr[(1-\varepsilon/3)\|y\|_2 \le \|B_T y\|_2 \le (1+\varepsilon/3)\|y\|_2 \text{ が不成立}]
$$
$$
\le \sum_{y\in N_T} \exp(-c_0 m) \qquad \text{(補題 4.1 の帰結)}
$$
$$
\le (4/\varepsilon)^{|T|}\exp(-c_0 m) \le (4/\varepsilon)^t \exp(-c_0 m)
$$

定理 5.2 の証明 (3)

$$
\Pr\left[\begin{array}{l}|\mathrm{supp}(x)| \le t \text{ を満たすある } x \in \mathbb{R}^n \text{ に対して} \\ (1-\varepsilon)\|x\|_2 \le \|Bx\|_2 \le (1+\varepsilon)\|x\|_2 \text{ が不成立}\end{array}\right]
$$
$$
\le \sum_{\substack{T\subseteq\{1,\dots,n\},\\ |T|\le t}} \Pr\left[\begin{array}{l}\text{ある } \tilde{x} \in \mathbb{R}^{|T|} \text{ に対して} \\ (1-\varepsilon)\|\tilde{x}\|_2 \le \|B_T\tilde{x}\|_2 \le (1+\varepsilon)\|\tilde{x}\|_2 \text{ が不成立}\end{array}\right]
$$
$$
\le \sum_{\substack{T\subseteq\{1,\dots,n\},\\ |T|\le t}} (4/\varepsilon)^t \exp(-c_0 m)
$$
$$
= \left(\sum_{i=1}^{t}\binom{n}{i}\right)(4/\varepsilon)^t\exp(-c_0 m)
$$
$$
\le (en/t)^t(4/\varepsilon)^t\exp(-c_0 m) = \exp(t\ln(en/t) + t\ln(4/\varepsilon) - c_0 m)
$$

事実（演習問題）

任意の $k \le n$ に対して，$\sum_{i=0}^{k}\binom{n}{i} \le (en/k)^k$

定理 5.2 の証明 (4)

$$
\exp(t\ln(en/t) + t\ln(4/\varepsilon) - c_0 m)
$$
$$
= \exp(3r\ln(en/3r) + 3r\ln(4/\varepsilon) - c_0 m) \qquad (\because\ t = 3r)
$$
$$
= \exp\left(3r\ln\left(\frac{4e}{3\varepsilon}\frac{n}{r}\right) - c_0 m\right)
$$
$$
< \exp\left(3r\ln\left(\frac{n}{r}\right)^{c'} - c_0 m\right) \qquad (c' > 0 \text{ はある定数})
$$
$$
= \exp(3c' r\ln(n/r) - c_0 m)
$$
$$
\le \exp(3c'\frac{1}{C}m - c_0 m) \qquad (\because\ m \ge Cr\ln(n/r))
$$
$$
= \exp\left(\left(\frac{3c'}{C} - c_0\right)m\right)
$$
$$
= \exp(-cm) \qquad (c > 0 \text{ はある定数})
$$

定理 5.2 の証明 (5)

ここまでのまとめ

$$\Pr\left[\begin{array}{l}|\mathrm{supp}(x)| \le 3r \text{ を満たすある } x \in \mathbb{R}^n \text{ に対して} \\ (1-\varepsilon)\|x\|_2 \le \|Bx\|_2 \le (1+\varepsilon)\|x\|_2 \text{ が不成立}\end{array}\right] < \exp(-cm)$$

したがって，

$$\Pr\left[\begin{array}{l}|\mathrm{supp}(x)| \le 3r \text{ を満たす任意の } x \in \mathbb{R}^n \text{ に対して} \\ (1-\varepsilon)\|x\|_2 \le \|Bx\|_2 \le (1+\varepsilon)\|x\|_2\end{array}\right] \ge 1 - \exp(-cm)$$

□

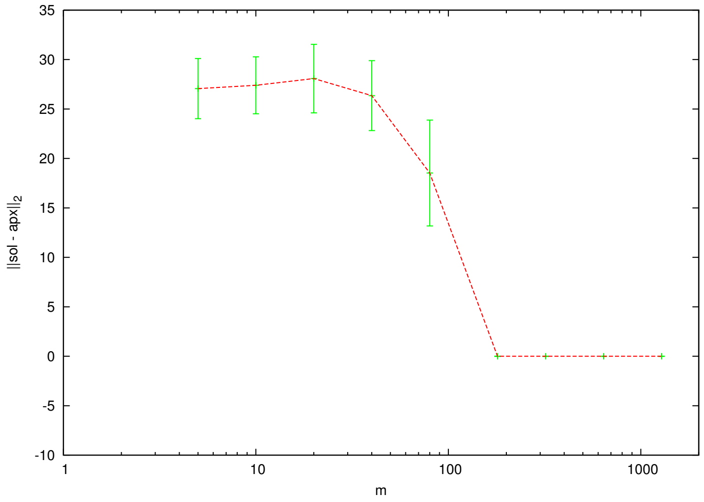
まとめ

- 線形方程式系の疎解探索
  - 多くの応用
  - NP 困難
- 基底追跡
  - 疎解探索に対するヒューリスティック解法
  - 線形計画問題
  - 係数行列が JL 補題のもの ⇒ 基底追跡の最適解が疎解

計算機実験

- $n = 2000, r = 20$：固定
- $m = 5, 10, 20, 40, 80, 160, 320, 640, 1280$：変化
- $A \in \mathbb{R}^{m \times n}$：各成分が独立に $\mathcal{N}(0,1)$ に従う
- $B = \frac{1}{\sqrt{m}} A$
- 疎解 $y \in \mathbb{R}^n$
  - 重複を許し，r 個の添字を一様ランダムに選択
  - 選んだ添字に対応する成分を $U(-10, 10)$ に従って選択
- $b = By$
- BP：$\inf\{\|x\|_1 \mid Bx = b\}$

計算機実験：プロット



m のとり得る値について

- 定理 5.2 では
$$m = \Omega(r \ln(n/r))$$
という下界があった (m が小さすぎてはいけない)
- これは次の意味でタイト

事実 (Donoho '06)

$m = o(r \ln(n/r))$ であるとき，定理 5.2 のように作成した B に対して基底追跡の唯一解は元の線形方程式系の疎解にならない

「中心対称近傍多面体 (centrally symmetric neighborly polytope)」に関する Linial, Novik ('06) の結果に基づく (Pfeifle '06 も参照)

$\rightsquigarrow$ 閾値現象の存在を示している

注意

- 「基底追跡」は疎解探索問題の緩和問題と見なすことができるが，昨日の講義で導入した意味で緩和問題になるわけではない
- 「疎解探索問題は組合せ的な問題なので NP 困難である」という言い方がよくされるが，正しくない
(組合せ的な問題でも簡単なものはある)

線形行列方程式の低階数解探索

線形行列方程式の低階数解探索

入力
- 線形関数 $\mathcal{A}\colon \mathbb{R}^{m\times n} \to \mathbb{R}^p$，ベクトル $b \in \mathbb{R}^p$，自然数 r

出力
- 方程式 $\mathcal{A}(X) = b$ の解 $X \in \mathbb{R}^{m\times n}$ で $\mathrm{rank}(X) \leq r$ を満たすもの

- 「線形方程式系の疎解探索」を特殊な場合として含む
- 多くの応用
- SDP を用いたヒューリスティック解法　(nuclear norm minimization)
  - RIP に類似した条件を $\mathcal{A}$ が満たす
  $\Rightarrow$ SDP ヒューリスティクスが低階数解を出力する
  - ランダムな線形関数 $\mathcal{A}$ は RIP を満たす

(Recht, Fazel, Parrilo '10)

Parrilo の発表 ('09) はよい情報源
http://www.mit.edu/~parrilo/pubs/talkfiles/ISMP2009.pdf

文献情報

今回の講義の内容は次の文献に基づく

- J. Matoušek. Basis pursuit and the Johnson–Lindenstrauss lemma (Lecture Notes). 2007.

それは次の論文に基づく

- R. Baraniuk, M. Davenport, R. DeVore, M. Wakin. A simple proof of the restricted isometry property for random matrices. Constructive Approximation 28 (2008) 253–263.

圧縮センシングに関する web 情報源 (膨大な関連文献情報)

- http://dsp.rice.edu/cs