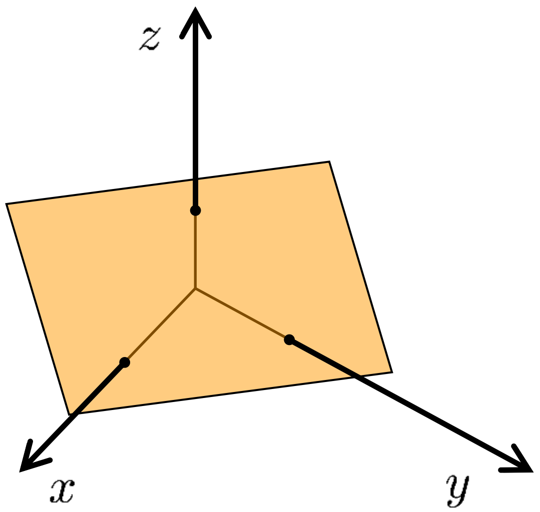
Day 2

方程式が足りない！どう解く？

Underdetermined problems
劣決定問題の近似解
最小長さ解
正則化

一次方程式

$$x \quad + \quad y \quad + \quad z \quad = \quad 1$$



連立一次方程式

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\left\{\begin{array}{ccccccc} x & + & y & + & z & = & 1 \\ x & + & 2y & + & 3z & = & 2 \\ x & & & + & z & = & -2 \end{array}\right.$$

問

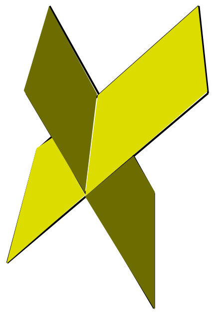
$$\left\{\begin{array}{ccccccc} x_1 & + & x_2 & + & x_3 & = & 1 \\ x_1 & + & 2x_2 & + & 3x_3 & = & 2 \\ x_1 & & & + & x_3 & = & -2 \end{array}\right.$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

劣決定の連立方程式

Underdetermined problem

方程式 m 本 $<$ 未知数 n 個



例：

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \end{cases}$$

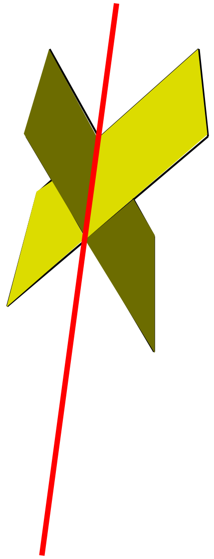
$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

劣決定の連立方程式

Underdetermined problem

方程式 m 本 ＜ 未知数 n 個

- 解が1つに決まらない
- 例では，直線上の点すべてが解の候補



例：

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \end{cases}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

連立一次方程式 ＝ 配合量を求める問題

混合物 $\boldsymbol{b}$ は,

- どの材料(＝ベクトル $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots, \boldsymbol{a}^{(n)} \in \mathbb{R}^m$) が使われているか？
- それぞれの配合量(＝ $x_1, x_2, \cdots, x_n$)はいくらか？

$$x_1 \boldsymbol{a}^{(1)} + x_2 \boldsymbol{a}^{(2)} + \cdots + x_n \boldsymbol{a}^{(n)} = \boldsymbol{b} \quad \Leftrightarrow \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\boldsymbol{A} = \left[\boldsymbol{a}^{(1)} \cdots \boldsymbol{a}^{(n)}\right] \in \mathbb{R}^{m \times n}$$

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

例:

$$x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\Leftrightarrow \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

優決定の連立方程式

Overdetermined problem

方程式 m 本 ＞ 未知数 n 個

例:

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \\ x_1 + x_3 = -2 \\ x_1 + 2x_2 + x_3 = 1 \end{cases}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}$$

優決定の連立方程式

Overdetermined problem

方程式 m 本　>　未知数 n 個

- 解が存在しない

例：

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \\ x_1 + x_3 = -2 \\ x_1 + 2x_2 + x_3 = 1 \end{cases}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}$$

正則化

Regularization

何らかの条件や情報，事前知識を使って，問題の解をひとつに限定すること.

色々な正則化の方法があります. いくつか後で紹介します.

- 解のℓ_2ノルムが小さいという条件： ℓ_2 正則化（ℓ_2 regularization）
  - リッジ回帰（ridge regression）, ティコノフ正則化（Tikhonov regularization）

- 解の非ゼロ成分の数が少ないという条件： スパース正則化（sparse regularization）
  ℓ_0最小化（ℓ_0 regularization）， ℓ_1最小化（ℓ_1 regularization）
  - ℓ_1正則化最小二乗法（ℓ_1-LS）, LASSO回帰（LASSO regression）
    - 基底追跡（basis pursuit; BP）
    - 基底追跡ノイズ除去（basis pursuit denoising; BPDN）
    - LASSO（least absolute shrinkage and selection operator）

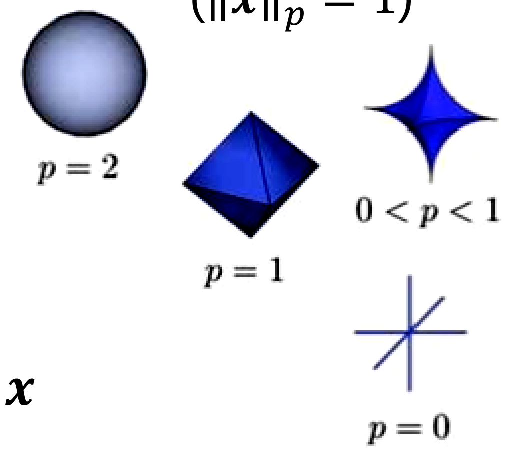
How to count nonzeros?

Sparsity-inducing norms ($p \le 1$)

ℓ_pノルム $\|\boldsymbol{x}\|_p^p = |x_1|^p + \cdots + |x_n|^p$

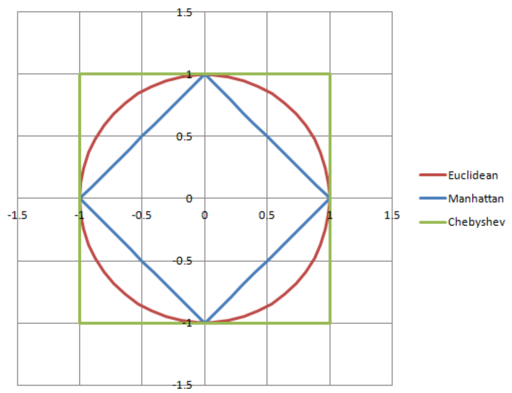
- $p = 2$ ℓ_2ノルム $\|\boldsymbol{x}\|_2 = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$
- $p = 1$ ℓ_1ノルム $\|\boldsymbol{x}\|_1 = |x_1| + \cdots + |x_n|$
- $p \to 0$ ℓ_0ノルム $\lim_{p\to 0} \|\boldsymbol{x}\|_p^p =$ **# of nonzero entries of** $\boldsymbol{x}$

ℓ_p-ball
($\|\boldsymbol{x}\|_p^p = 1$)




```
import numpy as np
# define a function of computing Lp norm to the power p
lpnormp = lambda x,p: sum(abs(x)**p)

# observe if the Lp norm converges to no. of nonzeros if p->0
x = np.array([1, 2, 0, 3, 0, 0, 4])
print(lpnormp(x, 1))        # 10
print(lpnormp(x, 0.5))      # 6.14626436994
print(lpnormp(x, 0.1))      # 4.33659499157
print(lpnormp(x, 0.01))     # 4.03196172178
```



最小長さ解

Minimum length solution

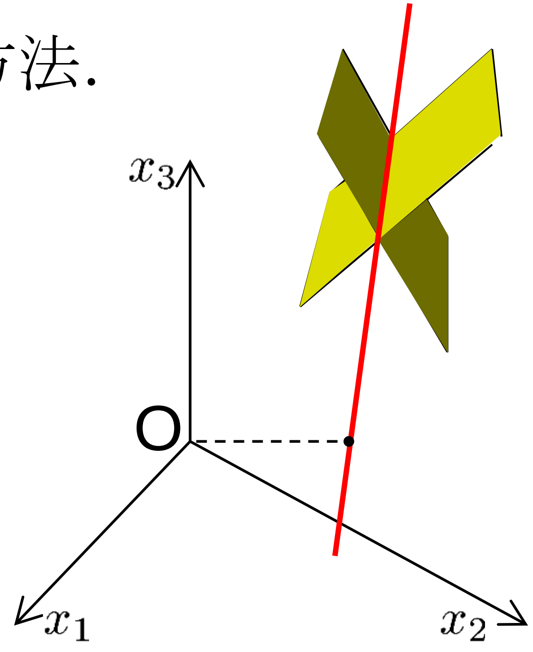
解の候補のうち, 原点に最も近いものを解に選ぶ方法.

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_2^2 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

問

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \end{cases}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

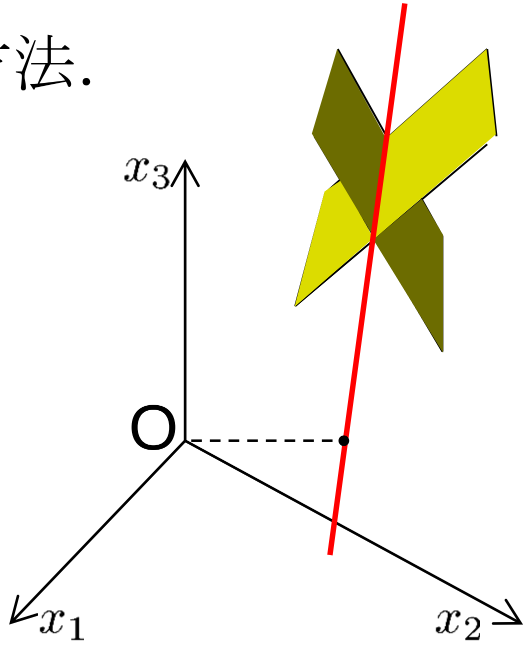


最小長さ解

Minimum length solution

解の候補のうち，原点に最も近いものを解に選ぶ方法.

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_2^2 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$



問

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \end{cases}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ 1-2t \\ t \end{bmatrix}$$

$$\|\boldsymbol{x}\|_2^2 = t^2 + (1-2t)^2 + t^2 = 6t^2 - 4t + 1$$

$$t = 1/3$$

最小長さ解

Minimum length solution

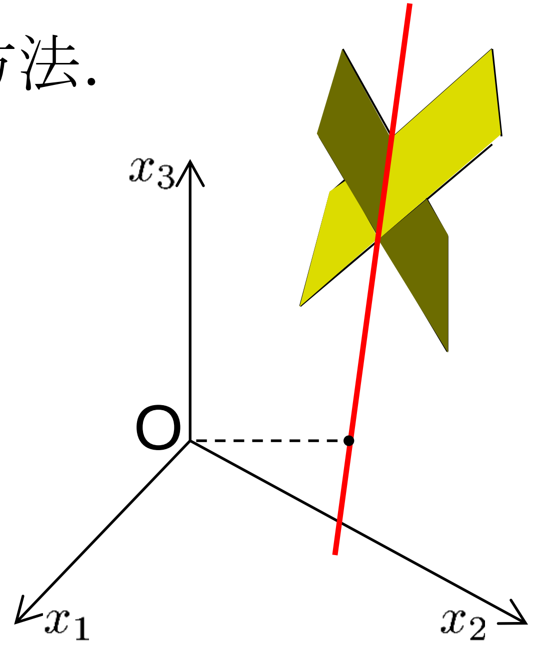
解の候補のうち, 原点に最も近いものを解に選ぶ方法.

$$\begin{aligned}\boldsymbol{x}^* &= \arg\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_2^2 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ &= \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}\boldsymbol{b}\end{aligned}$$

問

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_2 + 3x_3 = 2 \end{cases}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



ちょと
きゅうけい

多すぎるのは良くないことだ

More is less: when a linear system has a unique solution

劣決定問題のスパース解

味見してレシピを推定する

各食材を使用する
グラム数を求めたい

食材の線形結合

60×たまねぎ +30×にんじん +80×じゃがいも +40×豚肉

+2×香辛料A +・・・+0.1×香辛料X +10×油 +80×水 +120×ごはん

=

料理のスパースモデリング

$$m\left\{\underbrace{\begin{bmatrix} \boldsymbol{a}^{(1)} & \boldsymbol{a}^{(2)} & \boldsymbol{a}^{(3)} & \cdots & \boldsymbol{a}^{(n)} \end{bmatrix}}_{\boldsymbol{A}}\right. \underbrace{\begin{bmatrix} 60 \\ 30 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 40 \\ 0 \\ 0 \\ \vdots \end{bmatrix}}_{\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{b} \end{bmatrix}$$

食材集

料理を表すベクトル

各食材を使用する
グラム数（配合量）

各食材を
m 次元ベクトルで表す.

- どんな料理も，数種類の食材で作れる
- ベクトル $\boldsymbol{x}$ の成分はゼロだらけ（スパース）

連立一次方程式 ＝ 配合量を求める問題

混合物 $\boldsymbol{b}$ は,

- どの材料 (= ベクトル $\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots, \boldsymbol{a}^{(n)} \in \mathbb{R}^m$) が使われているか？
- それぞれの配合量 (= $x_1, x_2, \cdots, x_n$) はいくらか？

$$x_1\boldsymbol{a}^{(1)} + x_2\boldsymbol{a}^{(2)} + \cdots + x_n\boldsymbol{a}^{(n)} = \boldsymbol{b} \quad \Leftrightarrow \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\boldsymbol{A} = \left[\boldsymbol{a}^{(1)} \cdots \boldsymbol{a}^{(n)}\right] \in \mathbb{R}^{m\times n}$$

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

例：

$$x_1\begin{bmatrix}1\\1\end{bmatrix} + x_2\begin{bmatrix}1\\2\end{bmatrix} + x_3\begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix}$$

$$\Leftrightarrow \quad \begin{bmatrix}1 & 1 & 1\\1 & 2 & 3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix}$$

問：料理の食材を求めよ

$$
m\left\{\overbrace{\underbrace{\begin{bmatrix} \boldsymbol{a}^{(1)} & \boldsymbol{a}^{(2)} & \boldsymbol{a}^{(3)} & \cdots & \boldsymbol{a}^{(n)} \end{bmatrix}}_{\text{食材集}}}^{\boldsymbol{A}}\right. \overbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}^{\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{b} \end{bmatrix}
$$

トルコライスバーガー
が与えられたとき

各食材を使用する
グラム数を求めたい

各食材を
m 次元ベクトルで表す．

- 未知数 n 個　＞　方程式 m 本
- 解はひとつに定まらない．正則化が必要．
- スパースな解 $\boldsymbol{x}$ を選ぼう！

どの知識を選んで認識する？

誰？

Extended Yale Face Database B

どの知識を選んで認識する？

誰？

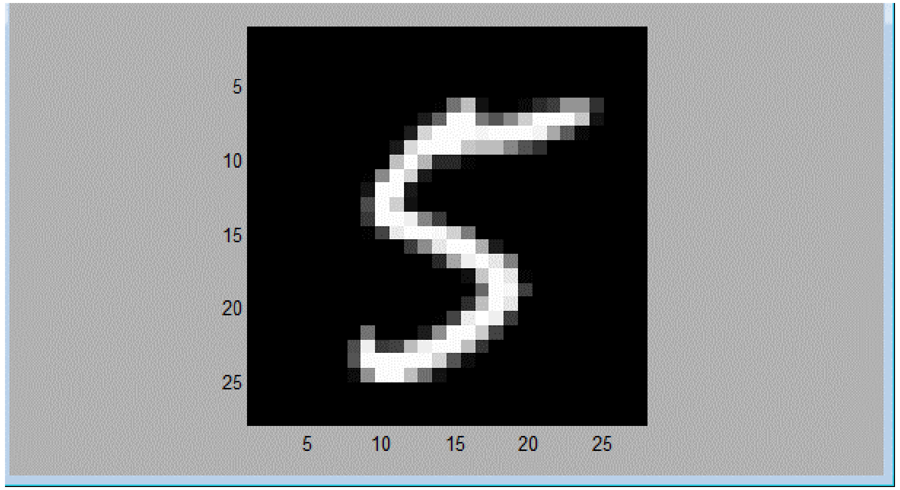
Extended Yale Face Database B

大量の知識の中から,
質問に対して簡潔な説明が作れる知識を選ぶ.

- 簡潔な説明＝厳選した少数の知識の組合せ

離散画像

コンピュータにとって，画像は数値の集まり



0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	4	114	192	18	0	0	11	43	61	148	148	49	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	29	99	252	252	120	85	137	206	252	253	252	201	21	0	0	0
0	0	0	0	0	0	0	0	0	0	0	31	213	252	252	252	245	252	252	252	244	170	91	11	0	0	0	0
0	0	0	0	0	0	0	0	0	0	27	218	252	252	252	199	190	189	136	84	49	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	194	253	182	42	42	7	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	4	139	253	212	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	22	252	252	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	75	252	208	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	57	252	252	239	133	59	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	4	68	235	253	252	249	211	124	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	62	141	232	253	253	175	29	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	41	168	239	253	231	130	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	17	200	252	233	7	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	104	252	247	63	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45	192	252	231	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	4	69	236	255	239	62	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	116	0	0	0	29	142	252	252	204	68	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	78	240	57	66	189	239	252	252	190	53	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	85	252	252	252	253	252	212	84	14	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	14	147	252	252	191	112	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

どの原料を選んで合成する？

質問画像

$\boldsymbol{b}$

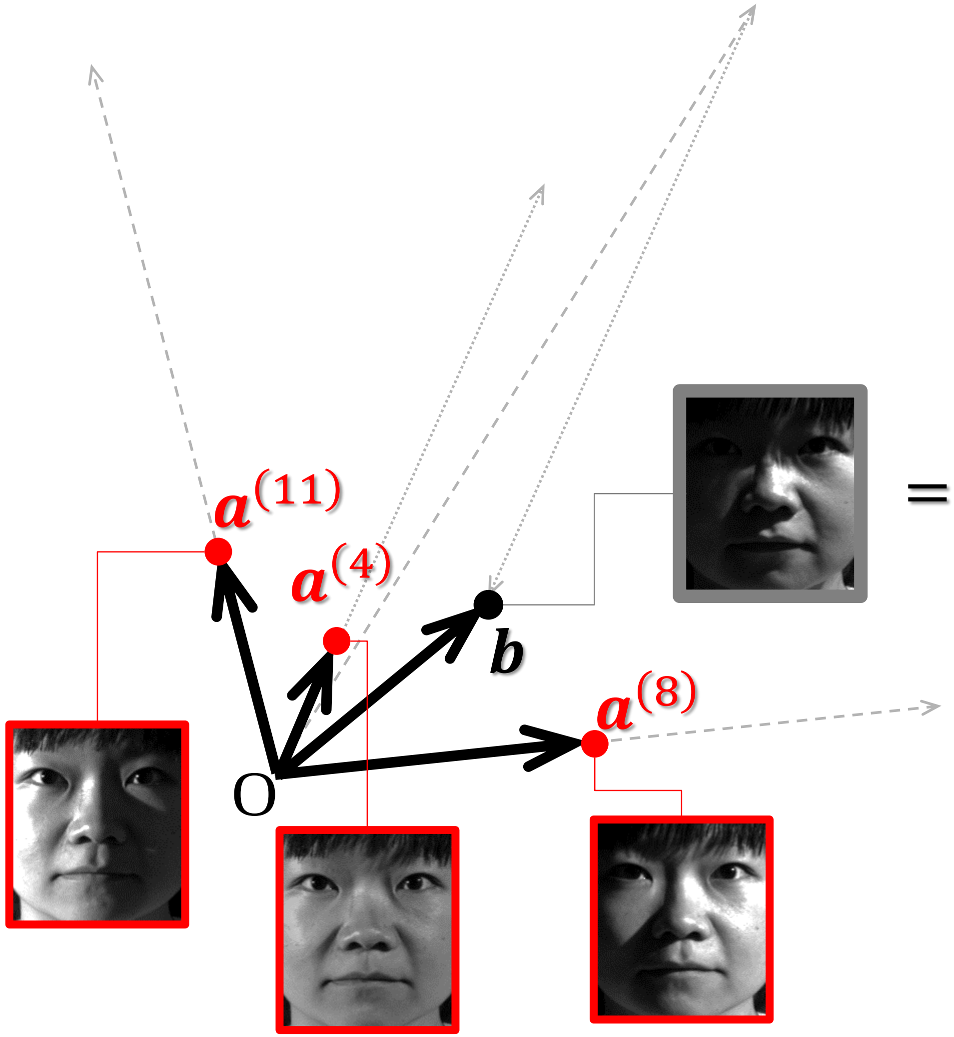
- 質問画像の合成

$$\boldsymbol{b} = x_1\boldsymbol{a}^{(1)} + x_2\boldsymbol{a}^{(2)} + \cdots + x_n\boldsymbol{a}^{(n)}$$

- 簡潔な説明＝使う画像の係数 x_j だけが非ゼロ

（疎, スパース(**sparse**)）

画像の和とスカラ倍



$$= -5\boldsymbol{a}^{(4)} + 2\boldsymbol{a}^{(8)} + 3\boldsymbol{a}^{(11)}$$

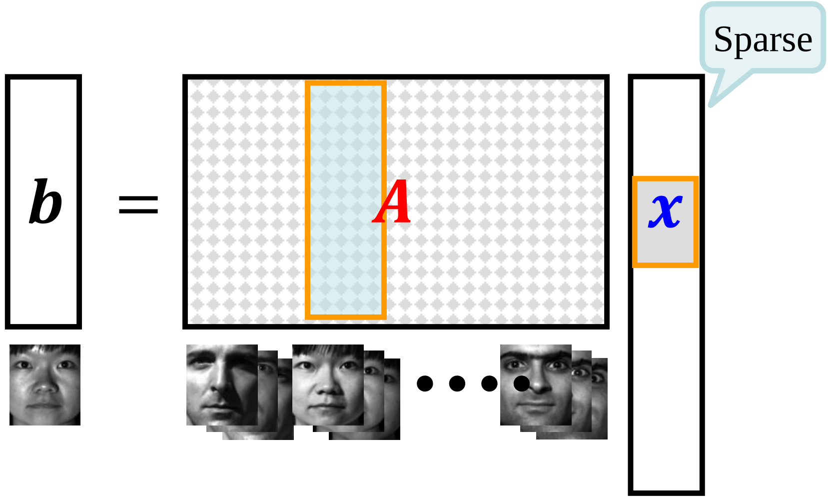
スパース表現に基づくパターン認識

[Wright+, 09]

データ $\boldsymbol{b}$ を簡潔に合成できる少数の原料を特定する.

$$\boldsymbol{b} = x_1\boldsymbol{a}^{(1)} + x_2\boldsymbol{a}^{(2)} + \cdots + x_n\boldsymbol{a}^{(n)}$$

$$= \begin{bmatrix} \boldsymbol{a}^{(1)} & \boldsymbol{a}^{(2)} & \cdots & \boldsymbol{a}^{(n)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$



どの原料を選んで合成する？

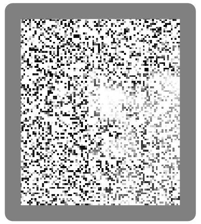
質問画像
$\boldsymbol{b}$

- 質問画像の合成

$$\boldsymbol{b} = x_1\boldsymbol{a}^{(1)} + x_2\boldsymbol{a}^{(2)} + \cdots + x_n\boldsymbol{a}^{(n)}$$

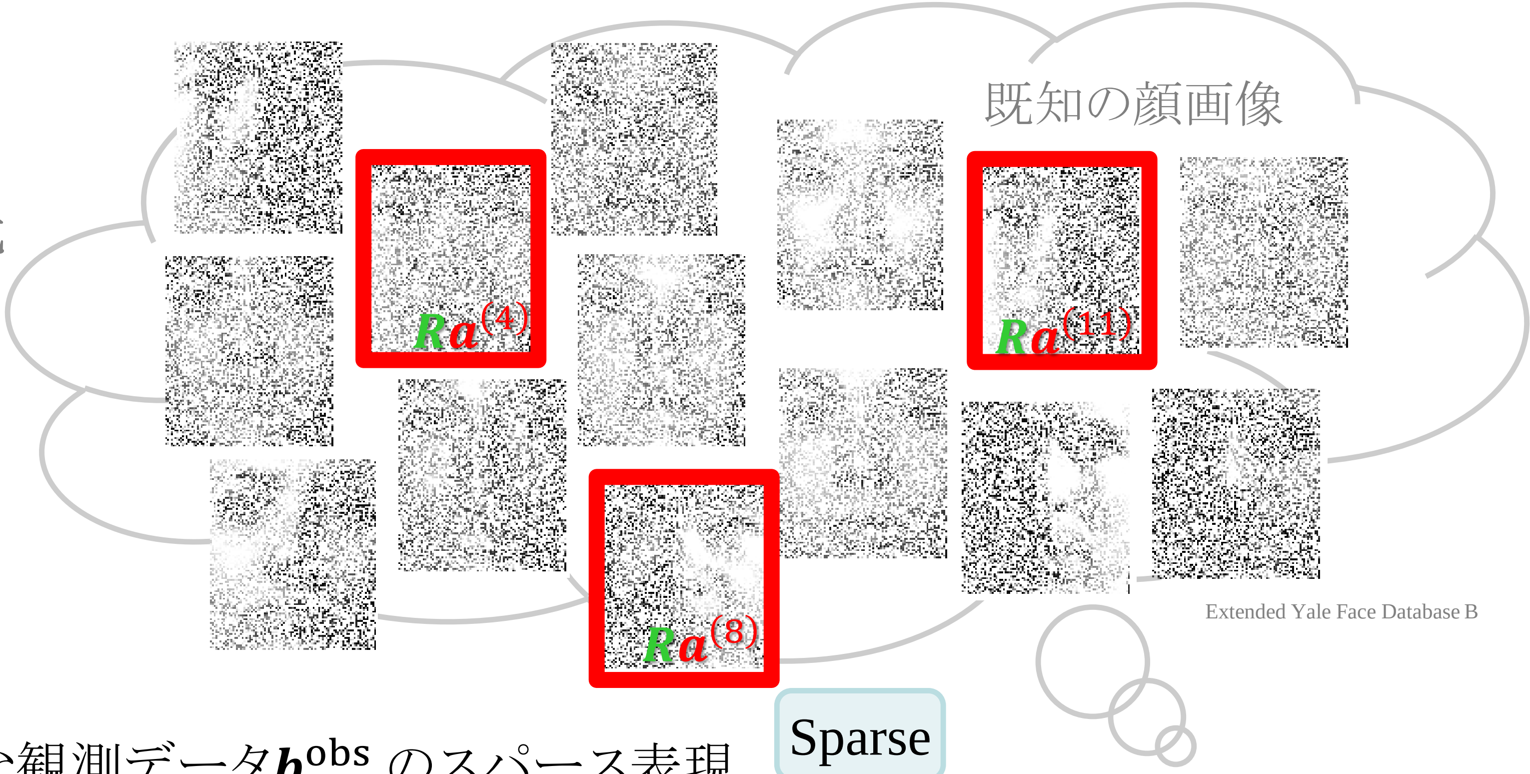
- 簡潔な説明＝使う画像の係数 x_j だけが非ゼロ
（疎, スパース(**sparse**)）

観測が不完全でも認識できる



ランダム標本化

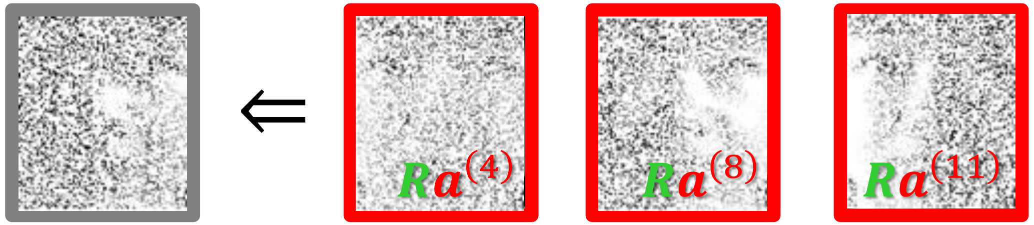
$\boldsymbol{b}^{\mathrm{obs}} = \boldsymbol{R}\boldsymbol{b}$



- 不完全な観測データ$\boldsymbol{b}^{\mathrm{obs}}$ のスパース表現

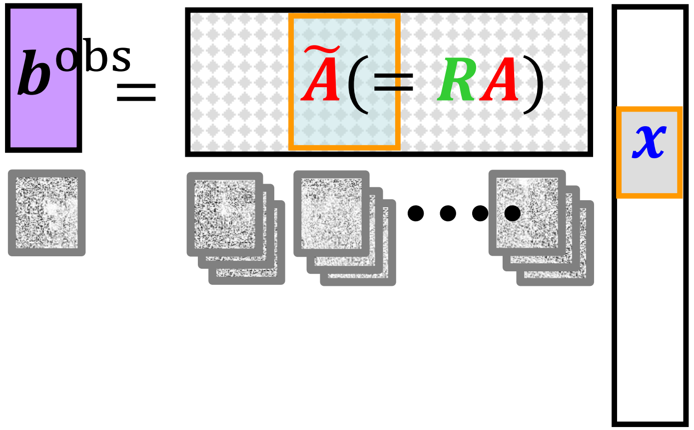
Sparse

$$\boldsymbol{b}^{\mathrm{obs}} = x_1 \boldsymbol{R}\boldsymbol{a}^{(1)} + \cdots + x_n \boldsymbol{R}\boldsymbol{a}^{(n)} = \tilde{\boldsymbol{A}}\boldsymbol{x} \quad (\tilde{\boldsymbol{A}} = \boldsymbol{R}\boldsymbol{A})$$



スパース表現に基づくパターン認識

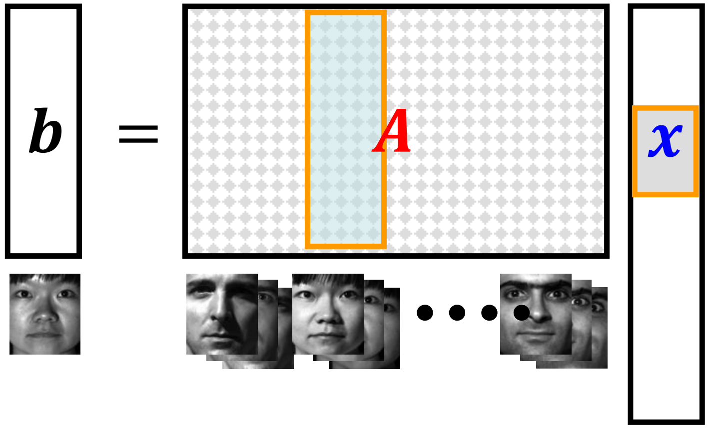
[Wright+, 09]



(1) 不完全なデータ$\boldsymbol{b}^{\mathrm{obs}}$
を上手に取得する.

(2) 適切な辞書（原料の集合）
を用いてスパース解 $\boldsymbol{x}$ を求める.

$$\underset{\boldsymbol{x}}{\text{Minimize}}\ \|\boldsymbol{x}\|_0\ \ \text{subject to}\ \ \widetilde{\boldsymbol{A}}\boldsymbol{x} \approx \widetilde{\boldsymbol{b}}$$



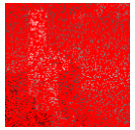
(3) 非ゼロ成分で判別できる.

圧縮センシング（compressed sensing）

[Donoho, 06] [Candes+, 06]

$\boldsymbol{b}^{\mathrm{obs}} = \tilde{A}(= RA)\,\boldsymbol{x}$

(1) 不完全なデータ$\boldsymbol{b}^{\mathrm{obs}}$
を上手に取得する.

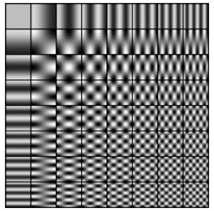


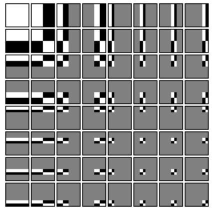
(2) 適切な辞書（原料の集合）
を用いてスパース解 $\boldsymbol{x}$ を求める.

$\underset{\boldsymbol{x}}{\text{Minimize}}\ \|\boldsymbol{x}\|_1$ subject to $\tilde{A}\boldsymbol{x} \approx \tilde{\boldsymbol{b}}$

(3) 完全なデータ $\boldsymbol{b}$ を再構成できる.

$\boldsymbol{b} = A\,\boldsymbol{x}$

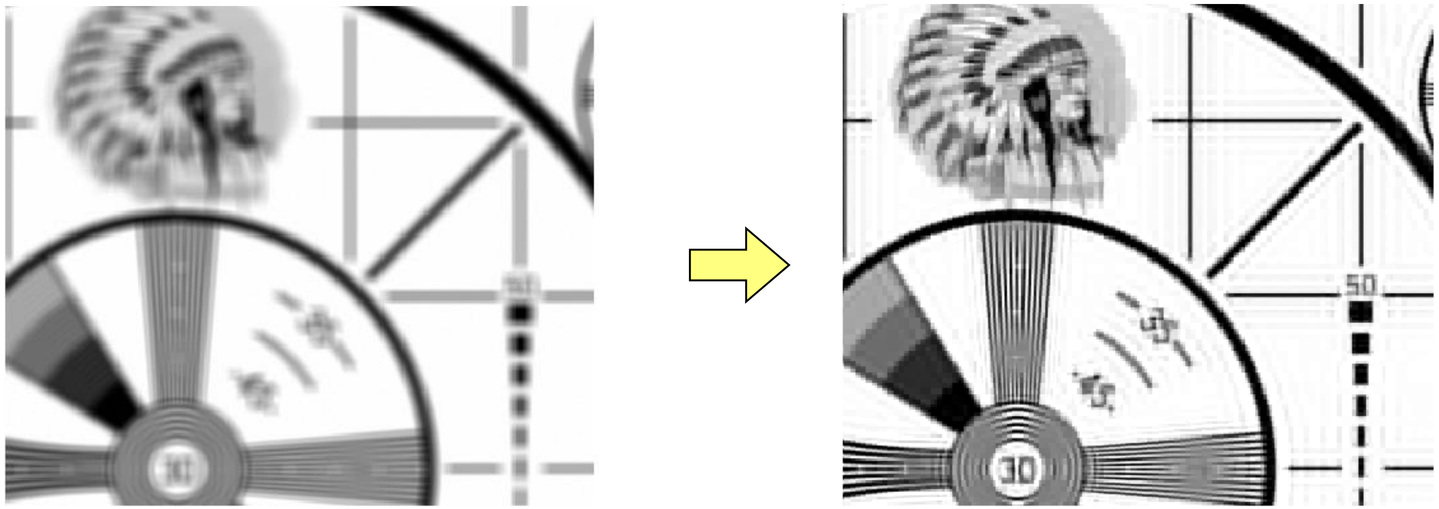




利点

- ✓ 計測の省力化
- ✓ データの高品位化
- ✓ 解析（内訳から本質を見抜く）
- ✓ ＋α

Deblurring / Superresolution



$\tilde{b} = RAx = \tilde{A}x$ $\qquad\qquad$ $b = Ax$

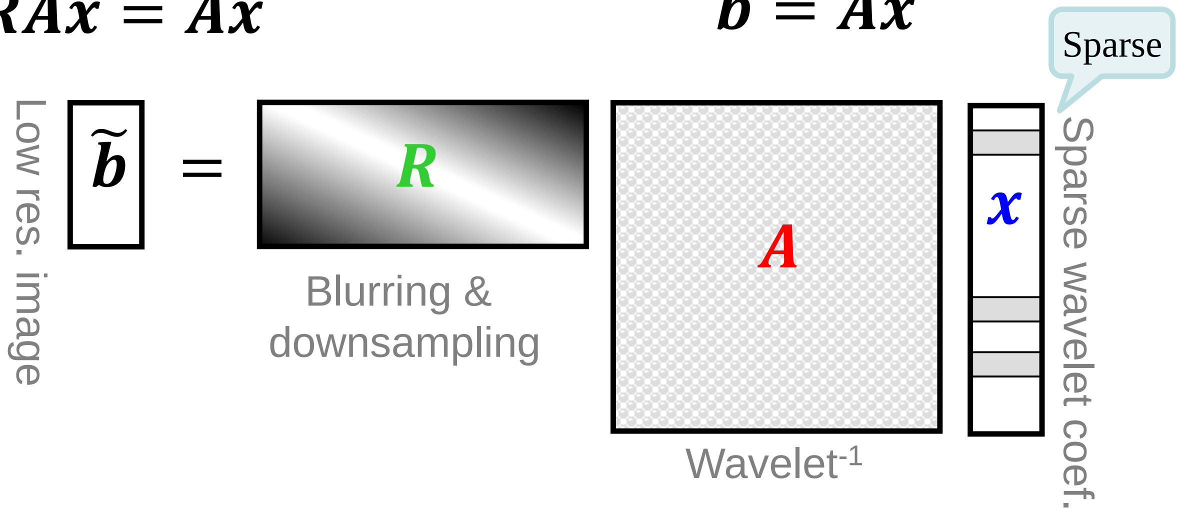
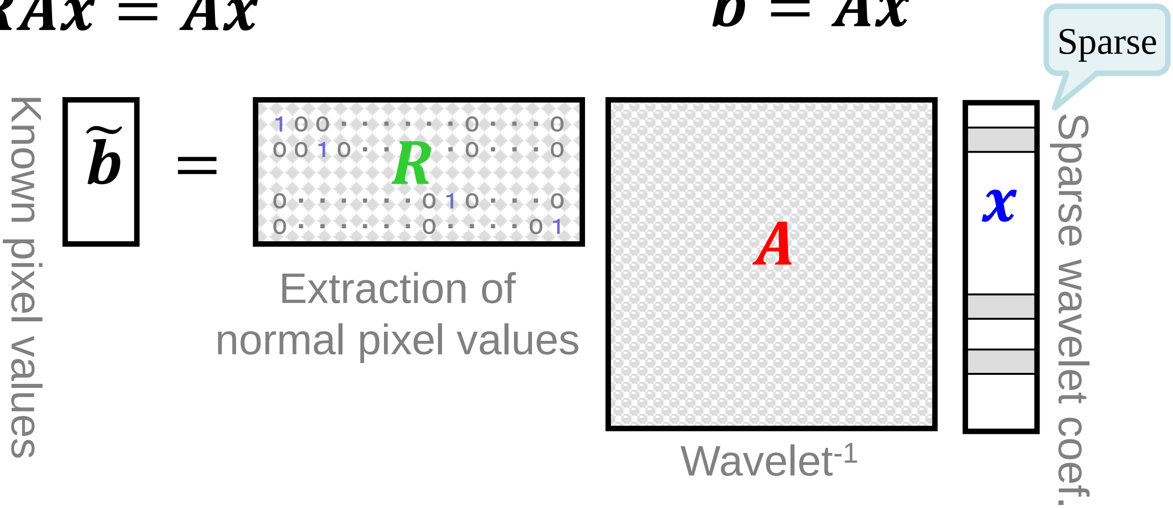


Image inpainting

$\tilde{\boldsymbol{b}} = \boldsymbol{RAx} = \tilde{\boldsymbol{A}}\boldsymbol{x}$

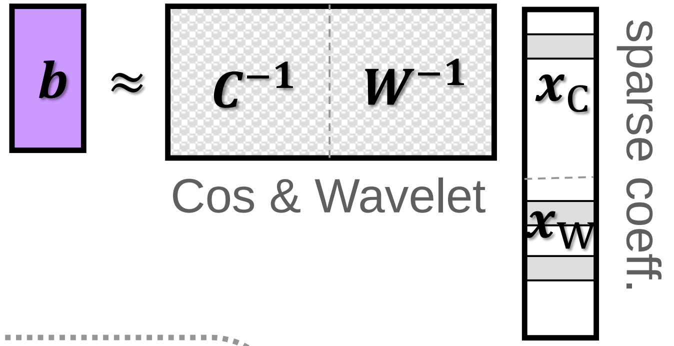
$\boldsymbol{b} = \boldsymbol{Ax}$

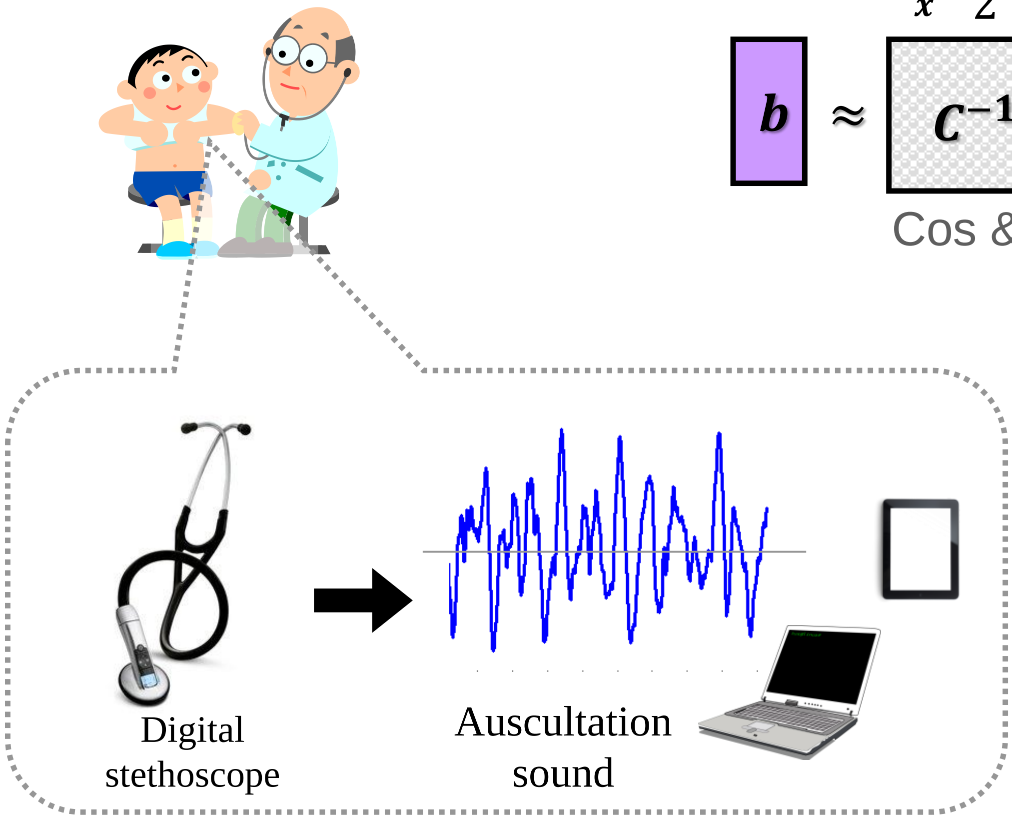


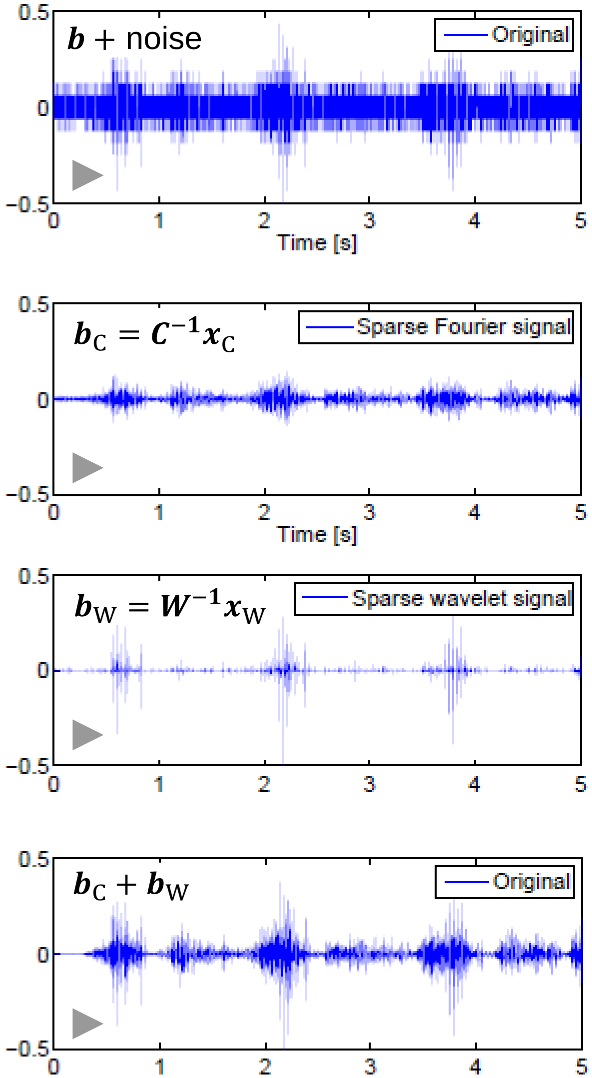
Lung sound separation (crackles)

[Sakai+11]

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_1$$







* Collaborating with the Second Department of Internal Medicine, Nagasaki University Hospital, Japan

劣決定問題のスパース正則化
Underdetermined Problem with Sparse Regularization

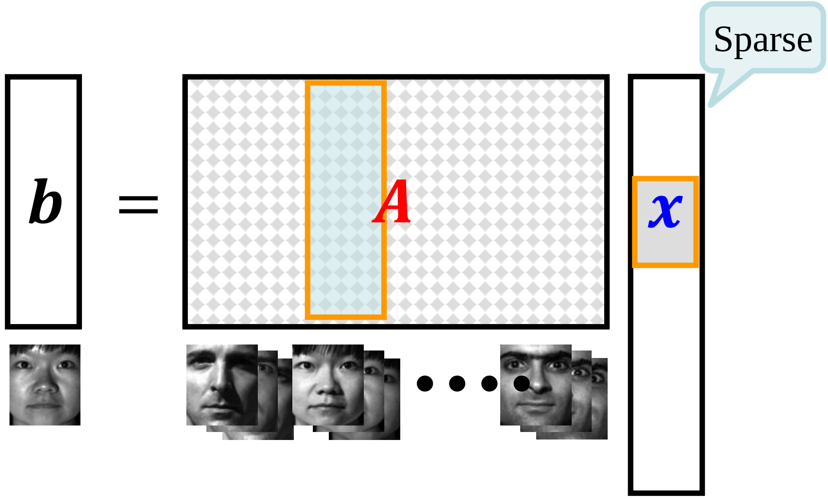
スパース表現に基づくパターン認識

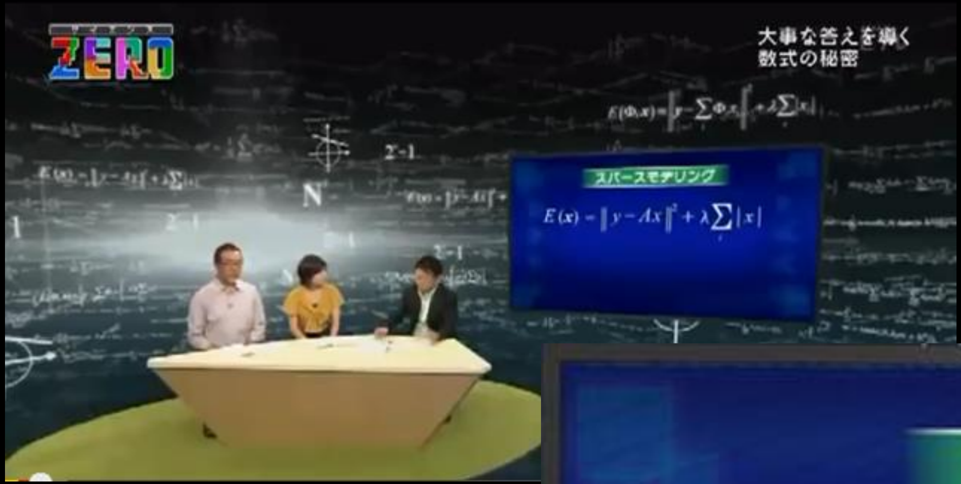
[Wright+, 09]

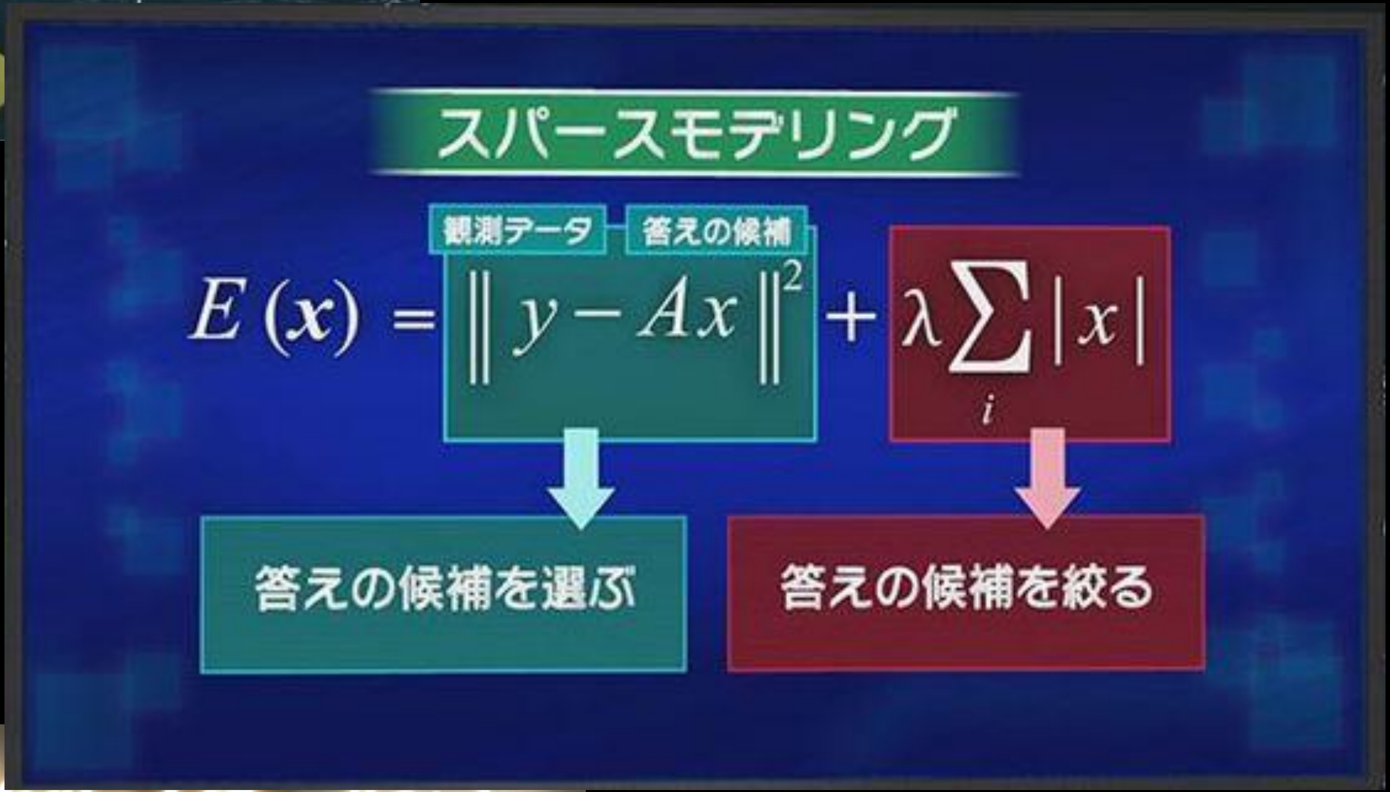
データ $\boldsymbol{b}$ を簡潔に合成できる少数の原料を特定する.

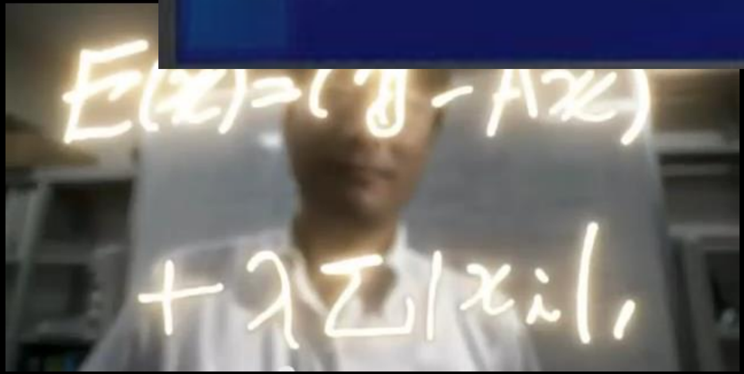
$$\boldsymbol{b} = x_1\boldsymbol{a}^{(1)} + x_2\boldsymbol{a}^{(2)} + \cdots + x_n\boldsymbol{a}^{(n)}$$

$$= \begin{bmatrix} \boldsymbol{a}^{(1)} & \boldsymbol{a}^{(2)} & \cdots & \boldsymbol{a}^{(n)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$




ZERO
大事な答えを導く
数式の秘密
スパースモデリング
$E(x) = \| y - Ax \|^2 + \lambda \sum_i |x|$


スパースモデリング
観測データ
答えの候補
$E(x) = \| y - Ax \|^2 + \lambda \sum_i |x|$
答えの候補を選ぶ
答えの候補を絞る


$E(x) = (y - Ax)$
$+ \lambda \sum |x_i|,$

Sparse solution

- Problem statement with ℓ_0 norm

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

"Find the sparse solution $\boldsymbol{x}^\star$, that satisfies $\boldsymbol{A}\boldsymbol{x}^\star = \boldsymbol{b}$, and has the smallest # of nonzeros".

- Issues
  - Uniqueness

    Is the sparse solution unique? Aren't there two or more solutions with the same # of nonzeros?
    - Under what conditions is the solution unique?
    - How sparse should the solution be?

  - Algorithm

    Can we efficiently find the sparse solution? Finding nonzeros seems to be combinatorial.

スパース解が一意に求る条件

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

スパースなベクトルに対して $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ が直交行列に似た性質をもっていれば,

- ただひとつのℓ_0最小解がある.
- 下記の ℓ_1最小解と等しい.

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

✓ 直交行列に似た性質

- 制限等長性（RIP） [Candes&Tao, 06]　　⇔　任意のkスパースベクトルに関するJL変換

$$(1-\delta_s)\|\boldsymbol{c}\|_2^2 \le \|\boldsymbol{A}\boldsymbol{c}\|_2^2 \le (1+\delta_s)\|\boldsymbol{c}\|_2^2 \quad \forall \boldsymbol{c} \in \{\, \boldsymbol{x} \mid \|\boldsymbol{x}\|_0 \le k\}$$

- 相互コヒーレンス [Donoho&Elad, 03] [Gribonval&Nielsen, 03]　⇔　原料の近似的な直交性

$$\|\boldsymbol{x}\|_0 \le \frac{1}{2}\left(1+\frac{1}{\mu}\right), \qquad \mu = \max_{i,j} \frac{|\boldsymbol{a}^{(i)\top}\boldsymbol{a}^{(j)}|}{\|\boldsymbol{a}^{(i)}\|\,\|\boldsymbol{a}^{(j)}\|} = \max_{i,j} |\cos\theta_{ij}|$$

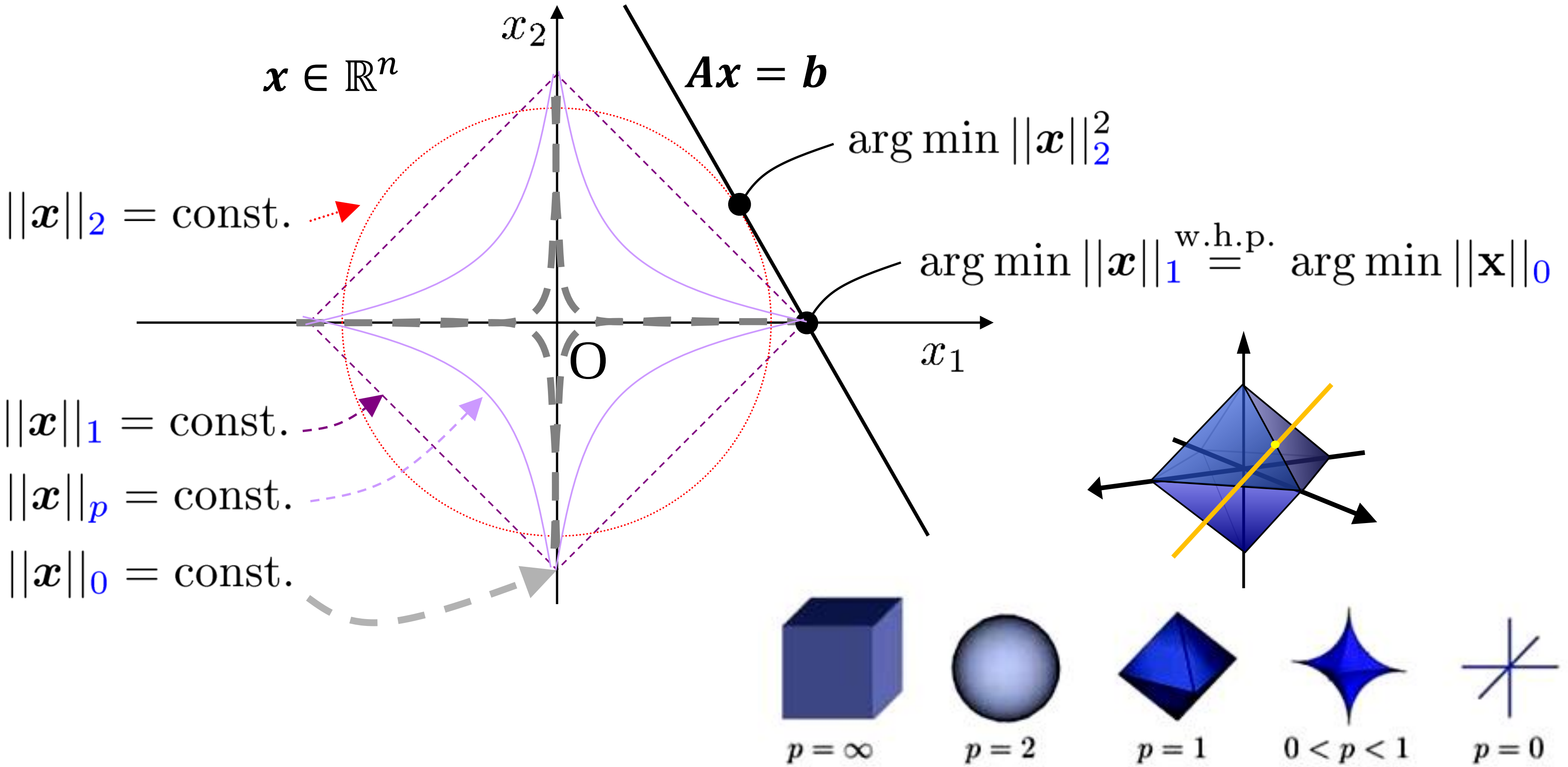
✓ $m = \Omega(k \log n)$ 個の計測データをもつ $\boldsymbol{b}$ から kスパース解（$\|\boldsymbol{x}\|_0 \le k$）が求まる.

✓ 制約条件を $\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2 \le \delta$ とすると, ノイズを含むデータ $\boldsymbol{b}$ をスパース近似できる.

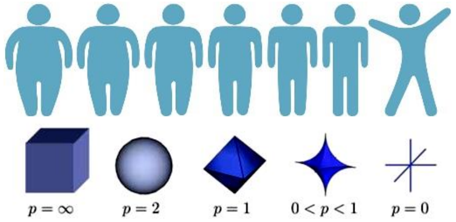
ℓ_0-ℓ_1 等価性の直観的な図解

$$\underset{\boldsymbol{x}}{\text{Minimize}}\ \|\boldsymbol{x}\|_p^p \ \text{ subject to }\ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$$

$$\|\boldsymbol{x}\|_p^p = |x_1|^p + \cdots + |x_n|^p$$



スパース解をもつ凸緩和問題



- BP: Basis pursuit

$$\underset{\boldsymbol{x}}{\text{Minimize}}\ \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$$

- BP$_\delta$: Constrained BP denoising

$$\underset{\boldsymbol{x}}{\text{Minimize}}\ \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2 \le \delta$$

- LASSO: least absolute shrinkage and selection operator

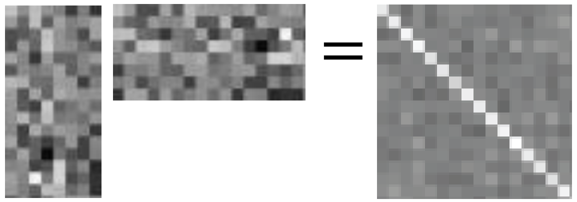
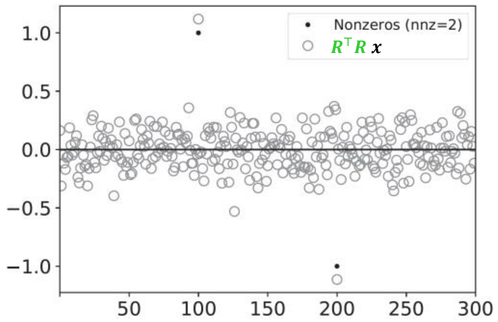
$$\underset{\boldsymbol{x}}{\text{Minimize}}\ \|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2 \quad \text{subject to} \quad \|\boldsymbol{x}\|_1 \le \tau$$

- LASSO regression / unconstrained BP denoising (BPDN) / ℓ_1-regularized least squares (ℓ_1-LS)

$$\underset{\boldsymbol{x}}{\text{Minimize}}\ \frac{1}{2}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2^2+\lambda\|\boldsymbol{x}\|_1$$

なぜランダム射影で維持できるのか？

- 直交行列$\boldsymbol{Q} \in \mathbb{R}^{m\times m}$
  - ✓ 正規直交性
    $$\boldsymbol{Q}^\top\boldsymbol{Q} = \boldsymbol{I}$$
  - ✓ ℓ_2ノルムの不変性
    $$\|\boldsymbol{Qx}\|_2^2 = \|\boldsymbol{x}\|_2^2$$
  - ✓ 可逆性
    $$\boldsymbol{Q}^\top(\boldsymbol{Qx}) = \boldsymbol{x}$$

- ランダム行列$\boldsymbol{R} \in \mathbb{R}^{m\times n}$
  （R_{ij} : iid Gaussian with mean zero and variance $1/m$）
  - ✓ 近似的な正規直交性
    （mutual incoherence）
    $$\boldsymbol{R}^\top \ \boldsymbol{R} \approx \boldsymbol{I}$$
    
  - ✓ 近似的なℓ_2ノルムの不変性（JL変換）
    $$(1-\varepsilon)\|\boldsymbol{x}\|_2^2 \le \|\boldsymbol{Rx}\|_2^2 \le (1+\varepsilon)\|\boldsymbol{x}\|_2^2$$
  - ✓ 近似的な可逆性
    （スパースベクトルに限る）
    $$\boldsymbol{x} = [0,0,\dots,0,1,0,\dots,0,-1,0,\dots,0]^\top$$
    $$\boldsymbol{x}^{\mathrm{est}} = \boldsymbol{R}^\top\boldsymbol{R}\ \boldsymbol{x}$$
    

今日は
ここまで