2013年2月18日 @九州大学

最適化ワークショップ：広がっていく最適化

行列の数値的同時ブロック対角化アルゴリズム

前原 貴憲[1,2]

1. 国立情報学研究所
2. JST, ERATO, 河原林巨大グラフプロジェクト

自己紹介

- 名前：前原 貴憲（まえはら たかのり）
- 略歴：

  –2012：東京大学 情報理工学系研究科 数理情報学専攻
  （指導教員：室田一雄教授）

  2012–現在： 国立情報学研究所 ERATO 特任研究員

- 興味：$\boxed{\text{最適化}}$・$\boxed{\text{数値計算}}$

行列の同時ブロック対角化問題

Given :

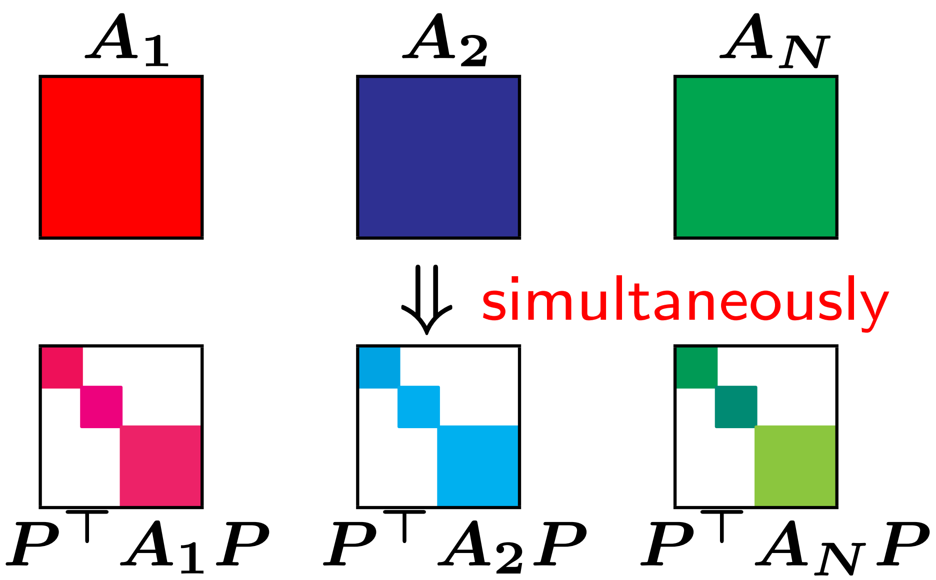
$n \times n$ 実正方行列 $A_1, \ldots, A_N$

Find :

直交行列 P s.t.

$P^\top A_1 P, \ldots, P^\top A_N P$: 同じ形のブロック対角行列

※ できるだけ細かい分解 を求めたい



本日の目標・発表のながれ

- 同時ブロック対角化問題
- モチベーション（応用先）
- 分解手法
- 課題・最近の進展

モチベーション（応用先）

同時ブロック対角化はどこであらわれるか

- 量子物理 [Wigner 1931–]

⋮

- 半正定値計画 [de Klerk-Dobre 2011] etc
- 完全正計画 [Arima-Kim-Kojima 2012] etc
- 非可換多項式計画 [Burgdorf-Klep-Povh 2011]
- 構造物の分岐経路追跡 [Aiura-Kakimura-Murota 2011]
- 独立成分分析 [Gutch-Krumsiek-Theis 2011]
- 動的ネットワークの安定性解析 [Irving-Sorrentino 2012]

⋮

古典的なテーマ，大規模問題の前処理として最近注目

量子物理（特に結晶物理） [Wigner 1931–]

一般に：物理量の計算＝固有値計算

$$Hx = \epsilon x$$

- $H = \alpha_1 H_1 + \cdots + \alpha_N H_N$
  物理パラメタの線型結合
- $H_1, \ldots, H_N$ 同時ブロック対角化
  $\Rightarrow$ パラメタに寄らない普遍的構造
- 結晶：高い対称性
  $\Longleftrightarrow$ 細かく分解可能（群表現論）

http://www.nasa.gov/centers/ames/multimedia/images/2006/crystal.html

動的ネットワークの安定性解析
[Irving-Sorrentino 2012]

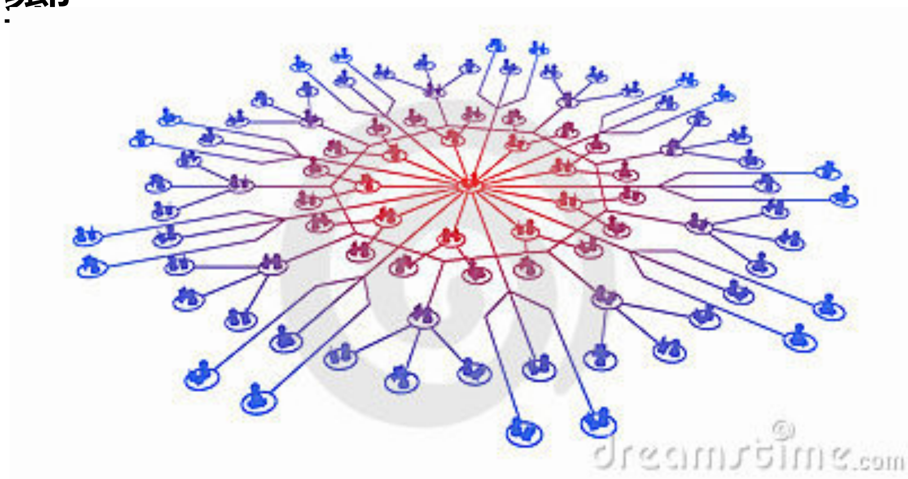
$$x(t) = A_1 x(t-1) + \cdots + A_N x(t-N)$$

- $x(t)$：時刻tで病気に感染している度合い
- A_k：k日前の感染者からの影響

例： A_1：第一近傍

A_2：第二近傍

$\vdots$



- 低次元なら豊富な安定性解析が使える

$\Rightarrow$ 低次元問題に分割＝同時ブロック対角化

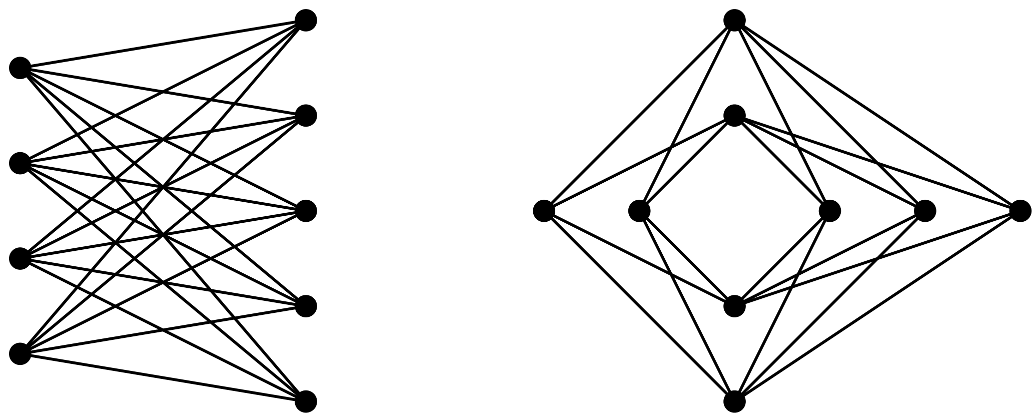
http://www.dreamstime.com/stock-image-network-infection-image11173181

半正定値計画問題

[Gatermann-Parrilo 2004], [Murota-Kanno-Kojima-Kojima 2010], ...

$$\begin{array}{ll} \text{minimize} & \langle \boldsymbol{C}, \boldsymbol{X} \rangle \\ \text{subject to} & \langle \boldsymbol{A}_i, \boldsymbol{X} \rangle = \boldsymbol{b}_i \ \ (\boldsymbol{i} = \boldsymbol{1}, \ldots, \boldsymbol{N}) \\ & \boldsymbol{X} \succeq \boldsymbol{O} \end{array}$$

- 係数行列の同時ブロック対角化
  $\Rightarrow$ 問題のサイズ・非線型性低減

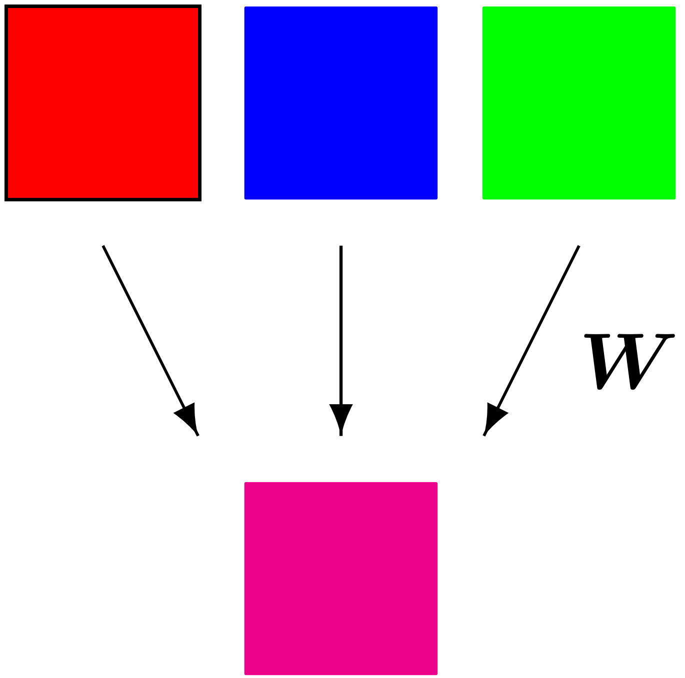


7 days $\longrightarrow$ 7 mins

交差数最小化 [de Klerk-Dobre-Pasechnik 2009]

独立成分分析
[Jutten-Herault 1985], [Cardoso-Souloniac 1993], ...

Y_1 Y_2 Y_ℓ

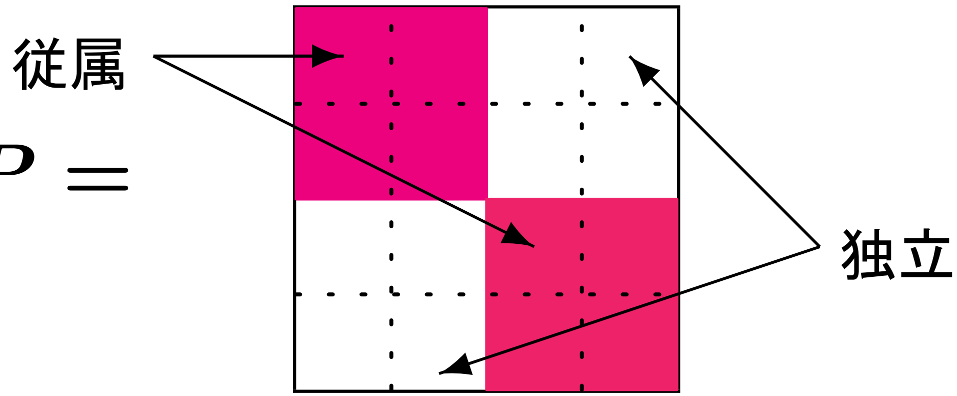


$X = WY$

Given: n次元観測信号 X

Find: 互いに独立な信号 $Y_1, \ldots, Y_\ell$

$$P^* C_\bullet P =$$



http://www.kindredcocktails.com/

再掲 同時ブロック対角化はどこであらわれるか

- 量子物理 [Wigner 1931–]

⋮

- 半正定値計画 [de Klerk-Dobre 2011] etc
- 完全正計画 [Arima-Kim-Kojima 2012] etc
- 非可換多項式計画 [Burgdorf-Klep-Povh 2011]
- 構造物の分岐経路追跡 [Aiura-Kakimura-Murota 2011]
- 独立成分分析 [Gutch-Krumsiek-Theis 2011]
- 動的ネットワークの安定性解析 [Irving-Sorrentino 2012]

⋮

古典的なテーマ，大規模問題の前処理として最近注目

分解手法

分解手法：２つの研究の流れ

- 数値線型計算：「SBD ＝ 固有値計算の一般化」
  - 数値的に安定な計算
  - 最細性の保証はないが，実用的に十分
  - 代数に詳しくない人でも理解しやすい

- 抽象代数：「SBD ＝ 群表現の分解」
  - 古典的な手法
  - 最細分解の保証
  - 近年数値計算の手法と合流

手法の一覧

numerical linear algebra | abstract algebra

old

one-by-one
[folklore -1820]

The recipe
[Schur 1905]

Jacobi-like
[Bunse-Gerstner, Byers, Mehrmann 1990]

JADE
[Cardoso-Souloumiacc 1993]

[Theis 2007]

MKKKM
[Murota-Kanno-Kojima-Kojima 2010]

MM
[Maehara-Murota 2012]

new

数值線型計算の系列

numerical linear algebra | abstract algebra

old

one-by-one
[folklore -1820]

The recipe
[Schur 1905]

Jacobi-like
[Bunse-Gerstner, Byers, Mehrmann 1990]

JADE
[Cardoso-Souloumiacc 1993]

[Theis 2007]

MKKKM
[Murota-Kanno-Kojima-Kojima 2010]

MM
[Maehara-Murota 2012]

new

one-by-one method [２入力・可換]

定理. ２つの可換な行列 $\boldsymbol{A}, \boldsymbol{X}$ は同時対角化可能

証明. wlog. $\boldsymbol{X} = \mathbf{diag}(\boldsymbol{x_1}, \ldots, \boldsymbol{x_n})$,

$$\boldsymbol{AX} - \boldsymbol{XA} = [(\boldsymbol{x_j} - \boldsymbol{x_i})\boldsymbol{a_{ij}}] = \boldsymbol{O}$$

$$\underset{\boldsymbol{X}}{\begin{bmatrix} x & & \\ & x & \\ & & y \end{bmatrix}} \qquad \underset{\boldsymbol{A}}{\left[\begin{array}{cc|c} * & * & \\ * & * & \\ \hline & & * \end{array}\right]}$$

「$\boldsymbol{X}$ を対角化して残った自由度で $\boldsymbol{A}$ を対角化する」

Jacobi-like method [２入力・可換] (Bunse-Gerstner, Byers, Mehrmann 1990)

one-by-one は誤差に弱い $\Rightarrow$ $\boxed{\text{最適化}}$ の考えで同時対角化

$$\mathbf{minimize}\ \ \left[\mathrm{off}(P^\top AP) + \mathrm{off}(P^\top BP)\right]$$

- Givens回転（$\mathbf{2} \times \mathbf{2}$ 回転）を繰り返して目的関数減らす（行列が１つなら固有値計算法の Jacobi 法と一致）
- 局所的には収束・大域的収束は不明

（元々の研究動機：歪ハミルトン構造をもつシステム解析）

JADE [多入力・可換]
(Cardoso, Souloumiac 1993)

$$\textbf{minimize}\ \left[\mathrm{off}(P^\top A_1 P) + \cdots + \mathrm{off}(P^\top A_N P)\right]$$

- Jacobi-like method の複数行列への拡張
- 信号処理（独立成分分析）で必要に

  1次元信号源 ＝ 同時対角化

- 実用的に，非常にうまくいっている
- 信号処理分野で標準的な手法の１つ

  （ソフトウェア公開 by Cardoso）

JADE' [多入力・非可換(ブロック対角化)] (Theis 2007)

$$\text{minimize}\ \left[\mathrm{off}(P^\top A_1 P) + \cdots + \mathrm{off}(P^\top A_N P)\right]$$

- 信号処理（独立空間分析）で必要に

  多次元信号源 = 同時ブロック対角化

- 目的関数の性質 :「最細分解」は局所最適解の１つ

  ⇒ 非対角でも構わず対角化用ルーチンを使う

うまくいかない例を発見 [Maehara–Gutch 2010]

- 高重複度問題
- 実型でない問題　　　$\cdots$　（代数的手法で「苦労した」例）

抽象代数の系列

numerical linear algebra

abstract algebra

old

one-by-one
[folklore -1820]

Schur lemma
[Schur 1905]

Jacobi-like
[Bunse-Gerstner, Byers, Mehrmann 1990]

JADE
[Cardoso-Souloumiacc 1993]

[Theis 2007]

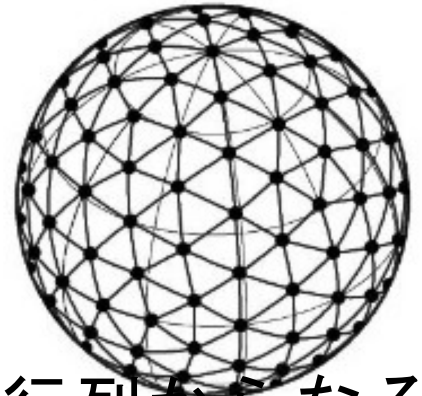
MKKKM
[Murota-Kanno-Kojima-Kojima 2010]

MM
[Maehara-Murota 2012]

new

対称性と同時ブロック対角化

行列 $\boldsymbol{A_1}, \ldots, \boldsymbol{A_N}$ が群 $\boldsymbol{G}$ の対称性をもつ：



$$\boldsymbol{Q}^\top \boldsymbol{A_j} \boldsymbol{Q} = \boldsymbol{A_j} \quad (\boldsymbol{Q} \in \boldsymbol{G})$$

$\boldsymbol{G}$：直交行列からなる群

定理．（Schurの補題）

- $\boldsymbol{G}$ を分解すると対応して $\boldsymbol{A_j}$ が分解される
- 逆に，$\boldsymbol{A_j}$ が分解できるときは非自明な $\boldsymbol{G}$ が存在

$\boldsymbol{G}$ の分解：表現の既約分解

- Schur lemma: 群論の標準的手法
- MKKKM, MM: one-by-oneの拡張（理論保証は代数的）

Schur lemma
(Schur 1905)

$$\boldsymbol{Q}^\top \boldsymbol{A}_j \boldsymbol{Q} = \boldsymbol{A}_j \quad (\boldsymbol{Q} \in \boldsymbol{G})$$

1. $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_N$ の対称性 $\boldsymbol{G}$ を発見
2. $\boldsymbol{G}$ の既約分解を計算
3. $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_N$ を分解

1930年付近で物理分野で流行りだす（Wignerが導入）
（cf. Heisenbergの行列力学 1925）

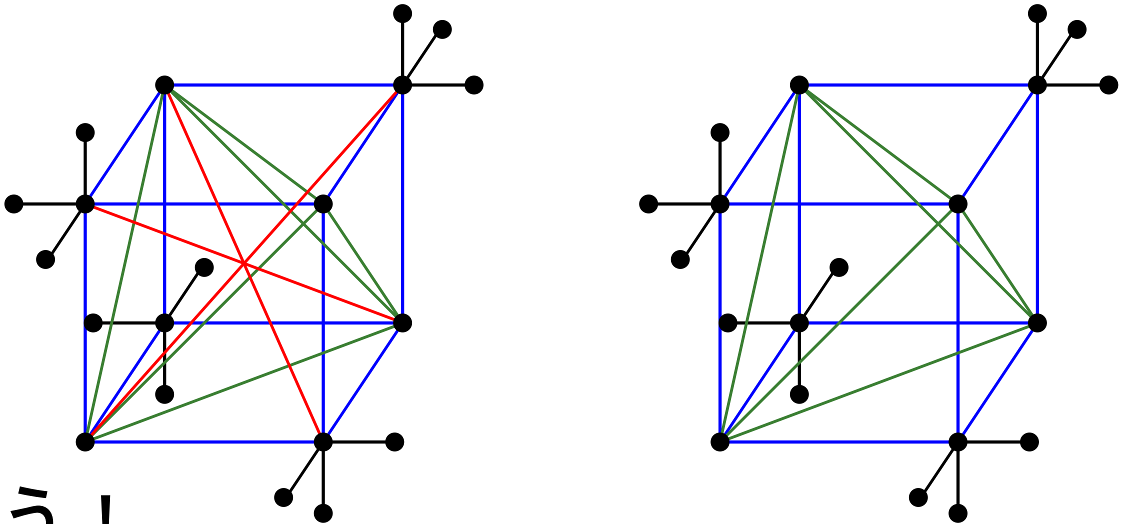
- 典型的な結晶の「既約分解」の表を作成
- 結晶に応じて表を組み合わせて既約分解（職人芸）

「対称性を人間が発見する」手法の限界

Q. これらの “対称性” は同じ？

$3V \times 3V$ 行列；$(3i+j, 3k+l)$ 成分＝

k 節点の第 l 次元が変化したときの i 節点の第 j 次元の変化



A. 違う！

- 幾何学的・組合せ的対称性は同じ
- 人間の手による「対称性発見」ではカバーしきれない

 $\Rightarrow$ 数値抽象代数的手法の導入へ

[Murota-Kanno-Kojima-Kojima 2010]

MKKKM
[Murota-Kanno-Kojima-Kojima 2010, Maehara-Murota 2011]

1. 行列を１つランダムサンプリングして対角化
   （$\boldsymbol{A_1}, \ldots, \boldsymbol{A_N}$ の和・積・スカラー倍・転置）
2. 対角化する行列を全体に適用
3. 残りの行列を残った自由度でブロック対角化

---

cf: one-by-one

1. １つの行列を対角化
2. 対角化する行列を全体に適用
3. 残りの行列を残った自由度で対角化

MKKKM抽象代数的理解
[Murota-Kanno-Kojima-Kojima 2010, Maehara-Murota 2011]

1. 行列を１つランダムサンプリングして対角化
   ＝行列∗代数の代表元を取得
2. 対角化する行列を全体に適用
   ＝単純成分分解
3. 残りの行列を残った自由度でブロック対角化
   ＝規約成分分解

---

cf: one-by-one

1. １つの行列を対角化
2. 対角化する行列を全体に適用
3. 残りの行列を残った自由度で対角化

MKKKM正当性議論の流れ

- $\boldsymbol{A_1, \ldots, A_N}$ のブロック対角化の構造
「Artin-Wedderburn構造定理」で完全に記述

定理. $\mathcal{T} := \langle \boldsymbol{A_1, \ldots, A_N} \rangle \simeq$

$$(\boldsymbol{M_{n_1}} \otimes \boldsymbol{I_{\mu_1}}) \oplus \cdots \oplus (\boldsymbol{M_{n_\ell}} \otimes \boldsymbol{I_{\mu_\ell}})$$

- $\mathcal{T}$ からランダムサンプル：
 ランダム行列の固有値は重複しない
 「固有値分解の一意性」より残る自由度： $\boldsymbol{M_n} \otimes \boldsymbol{I_\mu}$

※ one-by-one 同様，誤差に弱い

現行最良：MM法（アイデア）
[Maehara-Murota 2012]

one-by-oneの別方向への拡張；$\boldsymbol{X} = \mathrm{diag}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$,

$$\boldsymbol{AX} - \boldsymbol{XA} = [(\boldsymbol{x}_j - \boldsymbol{x}_i)\boldsymbol{a}_{ij}] = \boldsymbol{O}$$

$$\underset{\boldsymbol{X}}{\begin{bmatrix} x & & \\ & x & \\ & & y \end{bmatrix}} \qquad \underset{\boldsymbol{A}}{\left[\begin{array}{cc|c} * & * & \\ * & * & \\ \hline & & * \end{array}\right]}$$

- $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_N$と可換な対称$\boldsymbol{X}$を「見つけられたら」その固有値構造に応じて$\boldsymbol{A}_1, \ldots, \boldsymbol{A}_N$はブロック対角に
- 固有値がバラバラ→細かい分解：ランダムサンプリング

MM
[Maehara-Murota 2012]

1. $\boldsymbol{A_i X - X A_i = O}$ $(\boldsymbol{i = 1, \ldots, N})$をみたす対称行列をランダムサンプリング
2. $\boldsymbol{X}$ を対角化する直交行列を出力

- ランダムサンプル：固有値が最も分離する行列
  （厳密な証明：Artin-Wedderburn）

- 代数的には：交換子代数上の計算
  $\mathcal{T}' := \{\boldsymbol{X} \mid \boldsymbol{A_i X - X_i A = O}\ (\boldsymbol{i = 1, \ldots, N})\}$
  $\Rightarrow$ 「対称性の群」を数値的に計算
  ランダムサンプル：最も一般的な対称性
  固有値分解：群表現の既約分解

誤差制御版 MM
[Maehara-Murota 2012]

1. $\|\boldsymbol{A_i X} - \boldsymbol{X A_i}\| \le \boldsymbol{\epsilon}\ (\boldsymbol{i = 1, \ldots, N})$をみたす対称行列をランダムサンプリング
2. $\boldsymbol{X}$ を対角化する直交行列を出力

- 許容誤差に応じた分解（非対角ブロック成分に誤差）
- 非対角を丸めると$\boldsymbol{\epsilon}$近似ブロック対角行列
- 適切な$\boldsymbol{\epsilon}$：方程式系の最小特異値（正規誤差を仮定 $\Rightarrow$ 最小特異値 $\propto$ 誤差）

Error-Controlled Simultaneous Block-Diagonalization Algorithm

Problem

The **Simultaneous Block-Diagonalization Problem** is the following problem:

> **Given:** n x n real (or complex) matrices A_1, ..., A_N.
>
> **Find:** An orthogonal (or a unitary) matrix P such that P' A_1 P, ..., P' A_N P become block-diagonal matrices with a common diagonal structure (simultaneous block-diagonal form

The problem is widely studied in many areas such as physics, numerical computation, optimization, signal processing.

In many applications, the input matrices are contaminated with numerical errors (observation errors, approximation errors, etc.) So the error-controlled variant of the problem, **Error-Con Block-Diagonalization Problem**, is more useful in practice:

> **Given:** n x n real (or complex) matrices A_1, ..., A_N and *error-controlling parameter* $\varepsilon \geq 0$.
>
> **Find:** An orthogonal (or a unitary) matrix P such that P' A_1 P, ..., P' A_N P are *ε-approximated* simultaneous block-diagonal form, i.e., P' A_j P = (Block-Diagonal Matrix) + (O(

See [1] for more introduction and precise definition.

Algorithm

In [1], we have proposed the following simple algorithm, "MM-algorithm", for Error-Controlled Simultaneous Block-Diagonalization Problem.

1. Sample randomly from { X: n x n symmetric (or Hermitian) matrix | $|X A_j - A_j X| \leq \varepsilon$ (for all j) }.
2. Return an orthogonal (or a unitary) matrix P which diagonalizes X.

Step 1 is reduced to eigenvalue computation of a large matrix S (n^2 x n^2), and the Lanczos method can be applied.

Remark: The algorithm is a *dual* (in the sense of the *commutant* of the matrix *-algebra) version of the algorithm in [2,3], the so-called "MKKKM algorithm".

Source Code (in Matlab)

We have implemented the algorithm in Matlab.

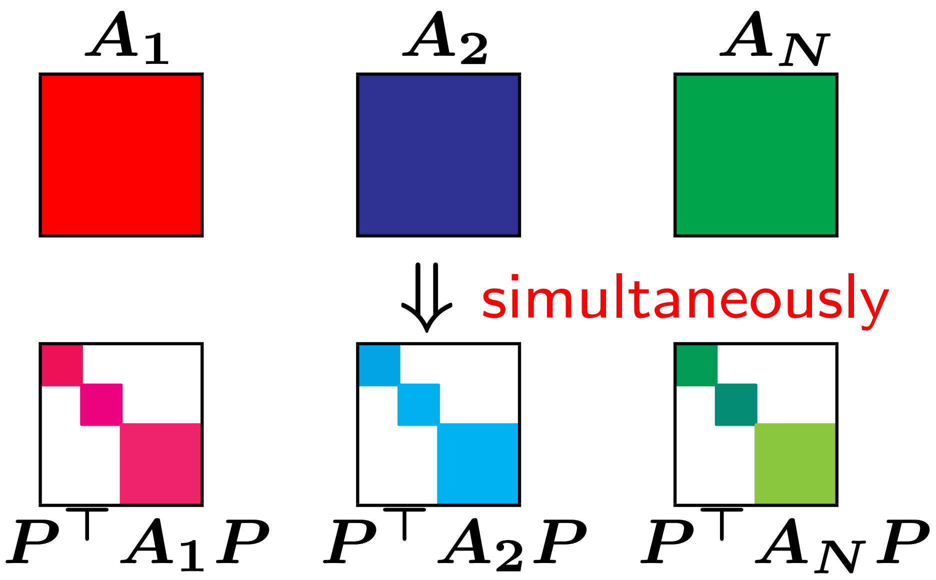
- commdec.m

The program is provided without warranty of any kind. You can use/modify/redistribute the program for any purpose.

http://www.misojiro.t.u-tokyo.ac.jp/~maehara/commdec/

ここまでのまとめ

行列の同時ブロック対角化問題



- 古典的問題，問題の大規模化に伴い近年活発に応用
- 手法：数値線型計算型，抽象代数型

  state of the art: 交換方程式の求解 [MM 2012]

課題・最近の進展

課題・最近の進展

- 効率的なアルゴリズム
  - 前処理（preconditioning）
  - 交互射影型アルゴリズム
- ランダム行列理論による評価
  - 定性的評価

ランダム行列理論による評価

$\boldsymbol{A}_i\boldsymbol{X} - \boldsymbol{X}\boldsymbol{A}_i = \boldsymbol{O}\ (i = 1, \ldots, N)$ をみたす
対称行列をランダムサンプリング

$\boldsymbol{X}$ の固有値分布，特に $|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_j|$ の分布を評価したい
$\Rightarrow$ ランダム行列の理論の結果が使えそう

定理 [Dyson index].
$\boldsymbol{n} \times \boldsymbol{n}$ 各成分標準正規のランダム対称行列の固有値密度

$$\propto \exp(-\boldsymbol{n} \sum \boldsymbol{\lambda}_i^2/4) \prod_{i<j} |\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_j|$$

（i.e., $\boldsymbol{\lambda}_i \simeq \boldsymbol{\lambda}_j$ になる確率は特に低い；反発効果）
（cf. ランダムエルミートにすると $|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_j|^2$）

ランダム行列理論による評価

$\mathcal{T}$：行列$*$代数，単純分解$\mathcal{T} = \mathcal{T}_1 \oplus \cdots \oplus \mathcal{T}_\ell$
$X \in \mathcal{T}$の固有値分布（with 適切なサンプリング）：

- 各ブロックのノルム $\propto n_k^2$
- 各ブロック内の分布 $\propto$ Dyson index
- 全体：これらの同時分布

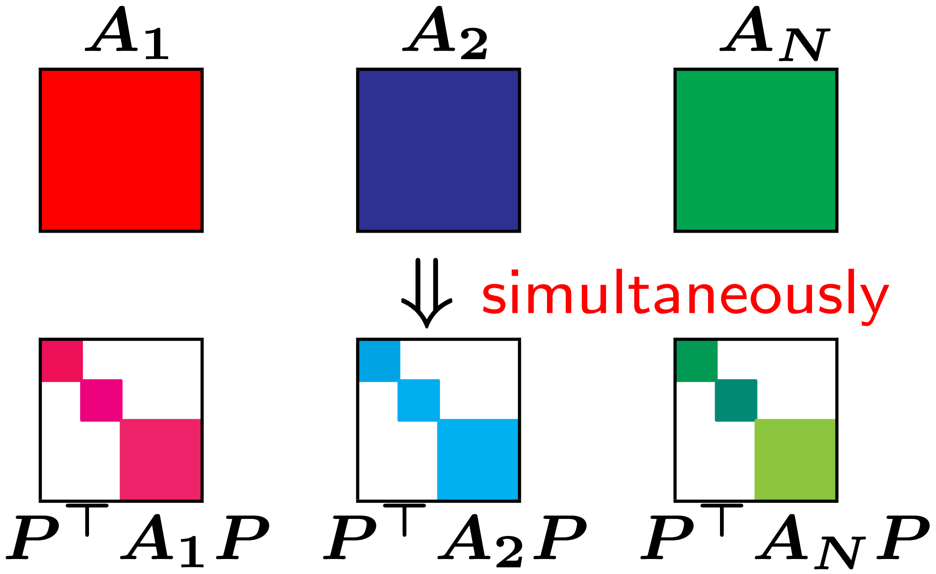
$\Rightarrow$

- ブロック数多は難しい（反発効果の無い固有値対増）
- ブロックサイズ比大は難しい（固有値がゼロ付近に密集）
- 複素数上で分解したほうが安定度が高い（反発効果の指数）

（ToDo: もう少し定量的な結果？）

まとめ

行列の同時ブロック対角化問題



- 古典的問題，問題の大規模化に伴い近年活発に応用
- 手法：数値線型計算型，抽象代数型

  state of the art: 交換方程式の特異値分解 [MM 2012]

- 課題・最近の進展

(以上)