

動的モード分解を活用した 高速将来予測アルゴリズム

千原直己^{†,‡}, 松原靖子[†], 藤原廉^{†,‡}, 櫻井保志[†]



† 大阪大学 産業科学研究所

‡ 大阪大学大学院 情報科学研究科

目次

- 研究背景
- モデル
- アルゴリズム
- 実験

目次

- **研究背景**
- モデル
- アルゴリズム
- 実験

研究背景

- 様々な事象により時系列データは生成されている

オンライン活動

具体例：ウェブ検索数, 口コミ

応用先：市場分析

Internet of Things (IoT)

具体例：振動, 加速度, 温度

応用先：自動運転

研究背景

- 様々な事象により時系列データは生成されている

オンライン活動

具体例：ウェブ検索数, 口コミ

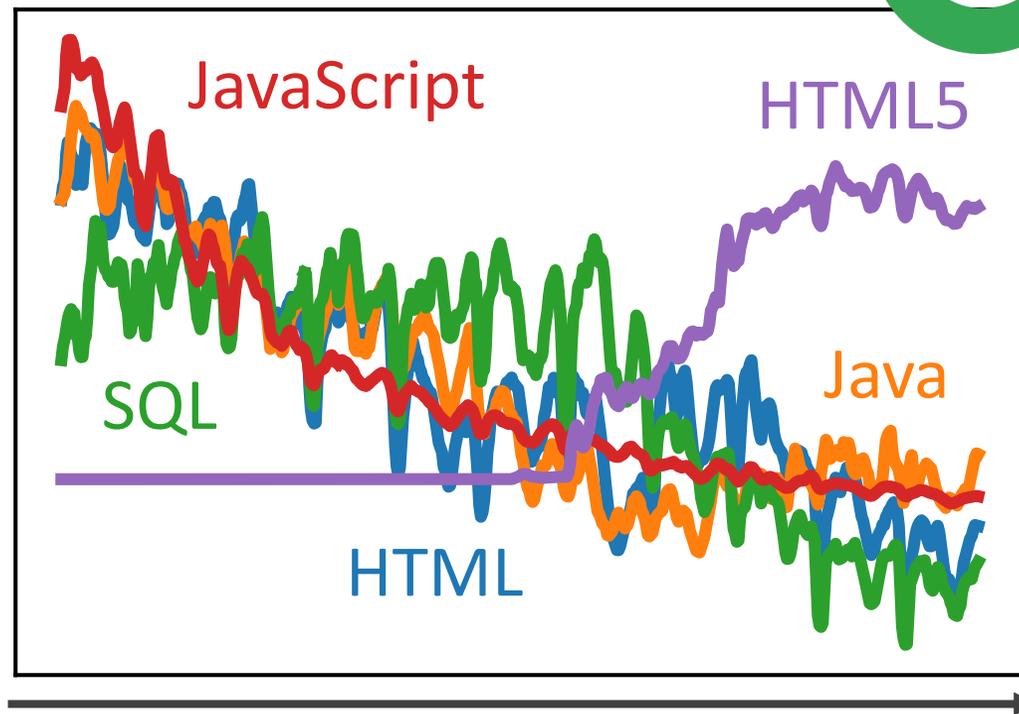
応用先：市場分析

Internet of Things (IoT)

具体例：振動, 加速度, 温度

応用先：自動運転

Google Trends



研究背景

- 様々な事象により時系列データは生成されている

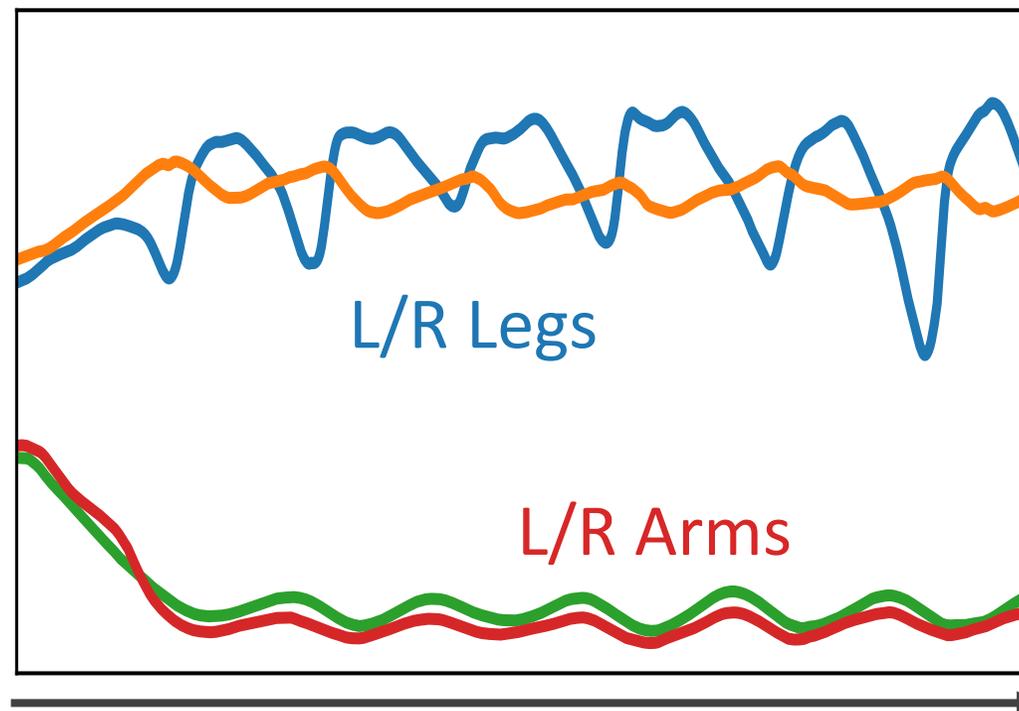
オンライン活動

具体例：ウェブ検索数, 口コミ
応用先：市場分析

Internet of Things (IoT)

具体例：振動, **加速度**, 温度
応用先：自動運転

Motion Sensors



研究背景

- 一般的な時系列データは複数の**潜在的パターン**を持つ

研究背景

- 一般的な時系列データは複数の潜在的パターンを持つ

Web データ

□ クリスマス

□ ブラックフライデー

研究背景

- 一般的な時系列データは複数の潜在的パターンを持つ

Web データ

□ クリスマス

□ ブラックフライデー

以後、このような潜在的なパターンを**レジーム**と呼ぶ

研究背景

- 一般的な時系列データは複数の潜在的パターンを持つ
- 高精度な予測のためには、これらを効果的に捉えることが必要不可欠である



上記を達成するために**動的モード分解 (DMD)** を活用する

- 非線形ダイナミクスの特徴を捉えることが可能

研究背景

- 一般的な時系列データは複数の潜在的パターンを持つ
- 高精度な予測のためには、これらを効果的に捉えることが必要不可欠である



上記を達成するために**動的モード分解 (DMD)** を活用する

- 非線形ダイナミクスの特性を捉えることが可能

- ❑ 指数関数的な増加/減少
- ❑ 周波数に基づく周期性

研究背景

- 一般的な時系列データは複数の潜在的パターンを持つ
- 高精度な予測のためには、これらを効果的に捉えることが必要不可欠である

動的モード分解により特徴的な**ダイナミクス**を捉え
リアルタイムな将来予測を実現する**ModeCast**を提案

研究背景: 問題定義

➤ Given

データストリーム

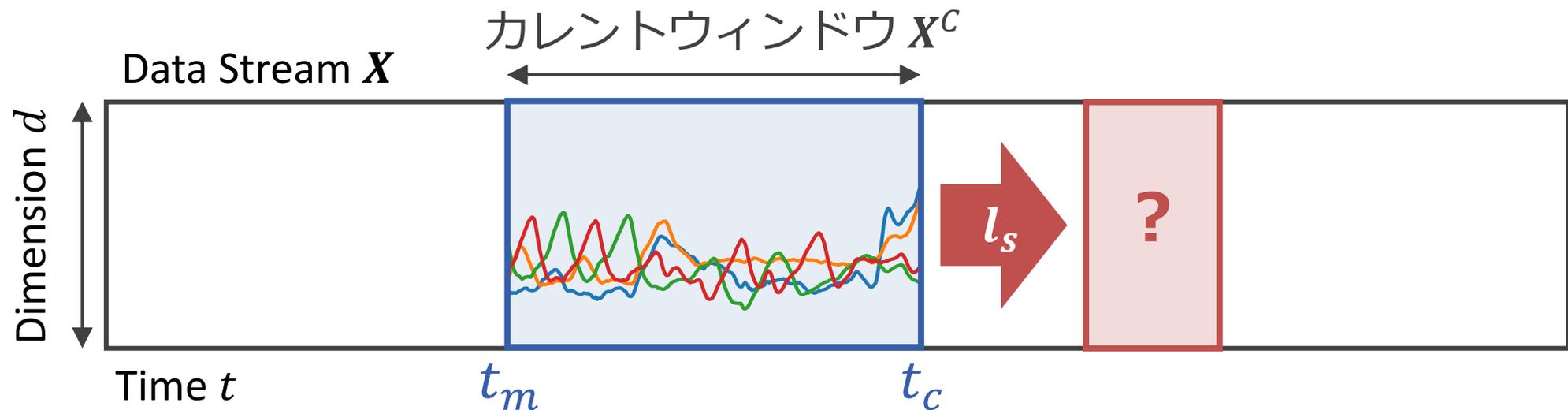
$$X = \{ \mathbf{x}(1), \dots, \mathbf{x}(t_c), \dots \}$$

t_c : 現在時刻

➤ Goal

任意の時刻における
 l_s ステップ先の予測

$$V^F = \{ \mathbf{v}(t_c + l_s), \dots \}$$



研究背景: 問題定義

➤ Given

データストリーム

$$X = \{ \mathbf{x}(1), \dots, \mathbf{x}(t_c), \dots \}$$

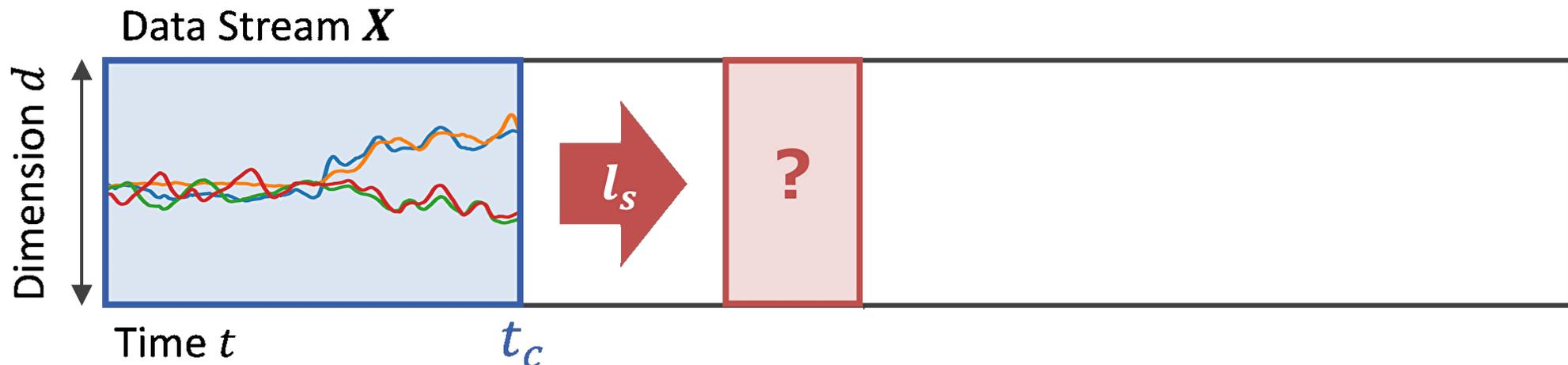
t_c : 現在時刻

➤ Goal

任意の時刻における

l_s ステップ先の予測

$$V^F = \{ \mathbf{v}(t_c + l_s), \dots \}$$



研究背景: 動的モード分解とは

- 非線形動的システムから潜在的な時間的ダイナミクスを抽出する数値解析手法

$$\frac{d}{dt} \mathbf{x} = \mathbf{f}(\mathbf{x})$$



$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t$$

研究背景: 動的モード分解とは

- 非線形動的システムから潜在的な時間的ダイナミクスを抽出する数値解析手法

$$\frac{d}{dt} \mathbf{x} = \mathbf{f}(\mathbf{x})$$

未知な非線形動的システム



$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t$$

線形動的システムで近似

研究背景: 動的モード分解とは

- 非線形動的システムから潜在的な時間的ダイナミクスを抽出する数値解析手法

Koopman Theory [Koopman 1931]
に基づいている

$$\frac{d}{dt} \mathbf{x} = \mathbf{f}(\mathbf{x})$$

未知な非線形動的システム



$$\mathbf{x}_{t+1} = \mathbf{A} \mathbf{x}_t$$

線形動的システムで近似

研究背景: 動的モード分解とは

- 非線形動的システムから潜在的な時間的ダイナミクスを抽出する数値解析手法
- 遷移行列 A は 2 種類の複素数範囲の行列で構成されている

$$A = \Phi \Lambda \Phi^\dagger$$

1. 動的モード Φ : k 次元空間の基底ベクトル
2. 固有値行列 Λ : 動的モードの時間発展

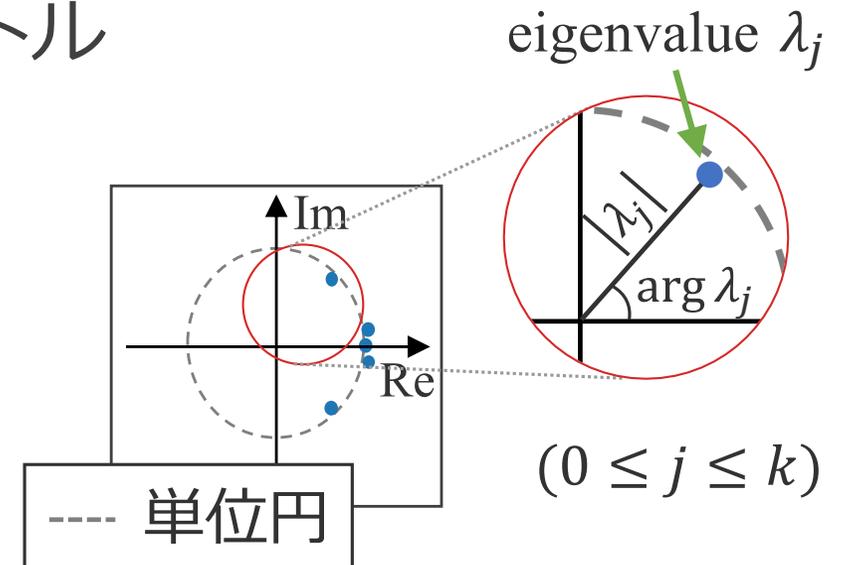
研究背景: 動的モード分解とは

- 非線形動的システムから潜在的な時間的ダイナミクスを抽出する数値解析手法
- 遷移行列 A は 2 種類の複素数範囲の行列で構成されている

$$A = \Phi \Lambda \Phi^\dagger$$

1. 動的モード Φ : k 次元空間の基底ベクトル
2. 固有値行列 Λ : 動的モードの時間発展

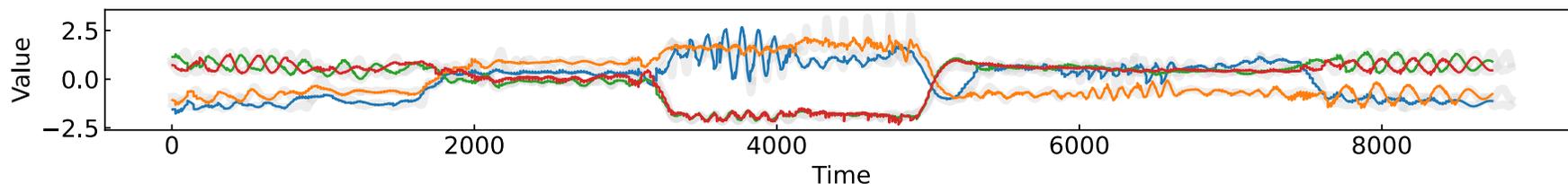
- $|\lambda_j|$: 減衰率
- $\arg(\lambda_j)$: 振動数



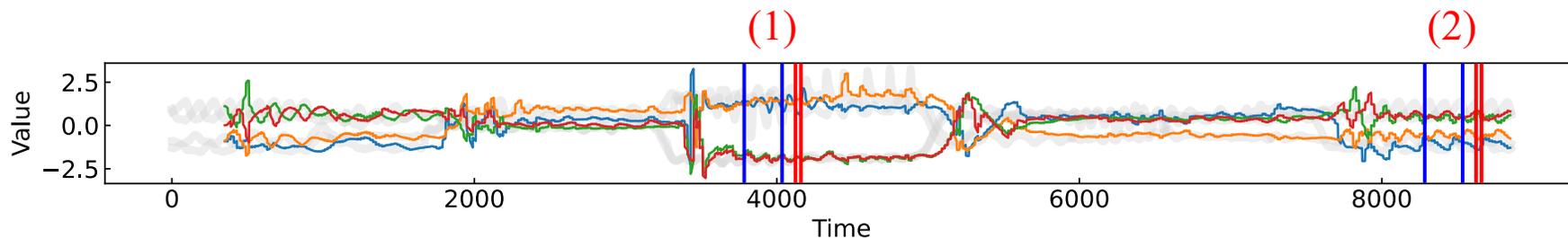
研究背景: 出力例

➤ センサデータストリームに対するリアルタイム予測

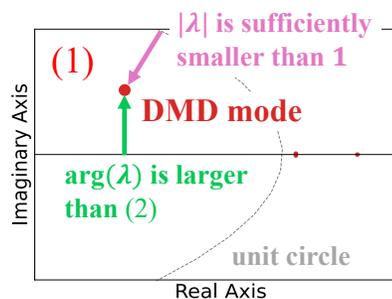
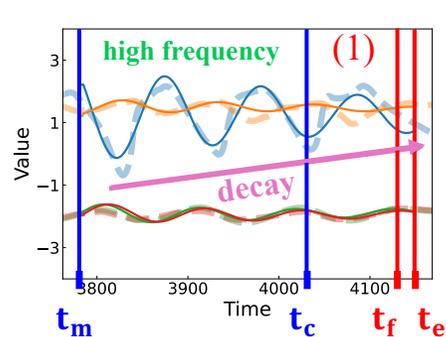
Fitting



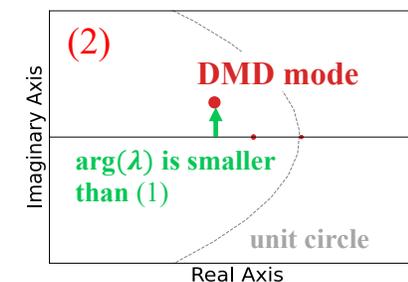
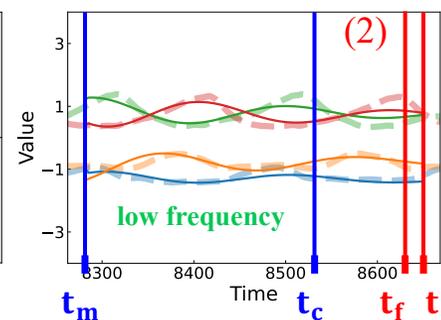
Forecast



Snapshots



$t_c = 4031$

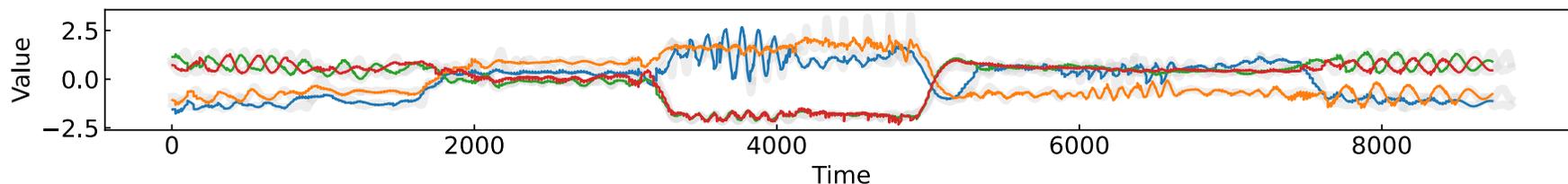


$t_c = 8531$

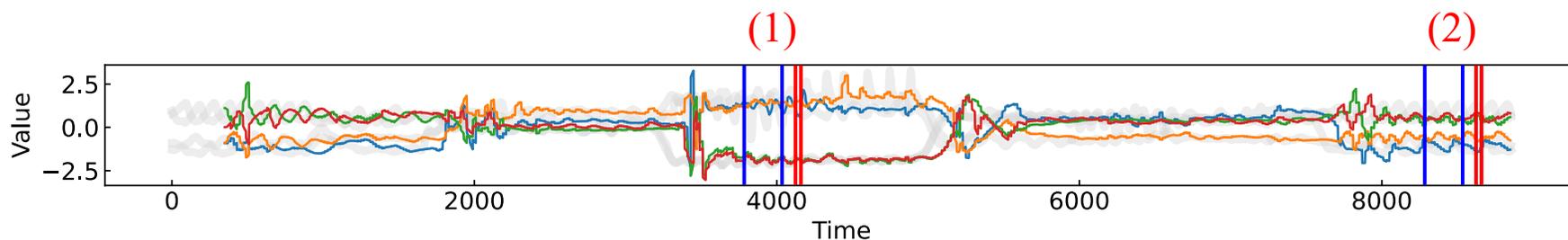
研究背景: 出力例

➤ センサデータストリームに対するリアルタイム予測

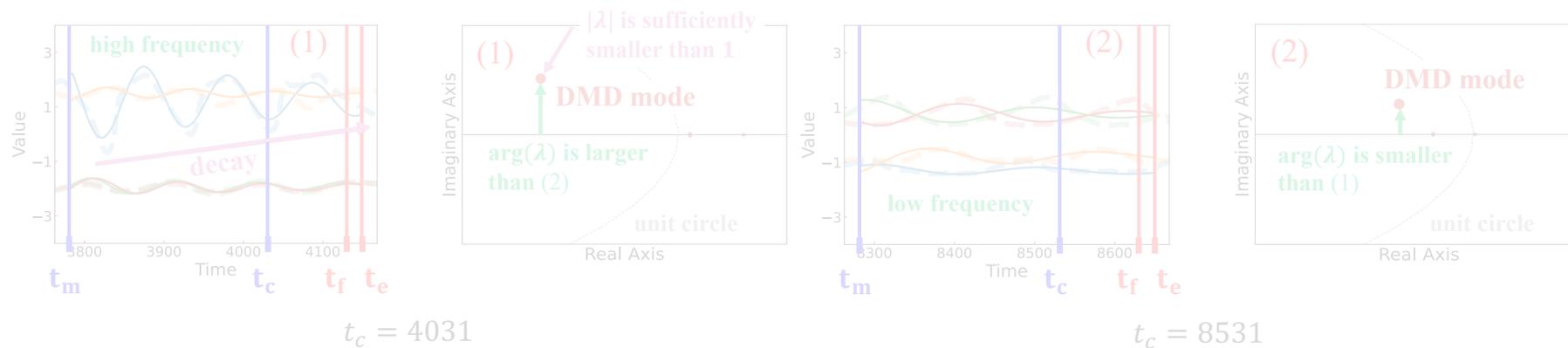
Fitting



Forecast



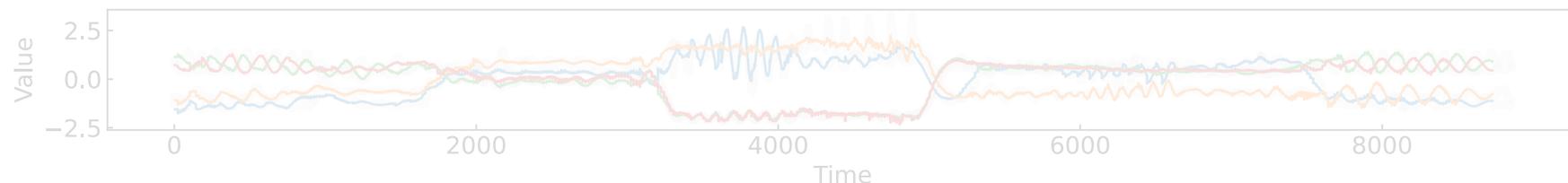
Snapshots



研究背景: 出力例

➤ センサデータストリームに対するリアルタイム予測

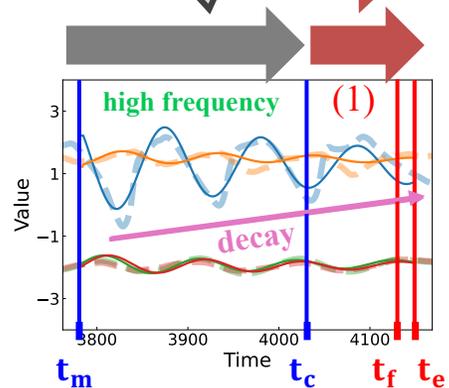
Fitting



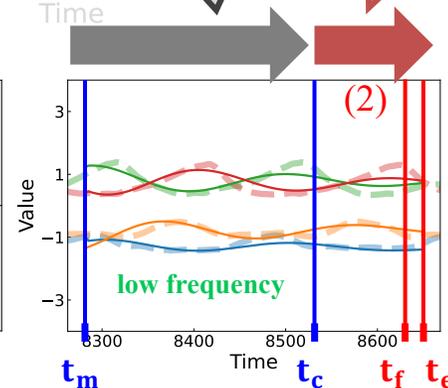
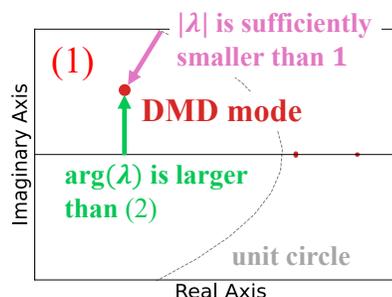
Forecast



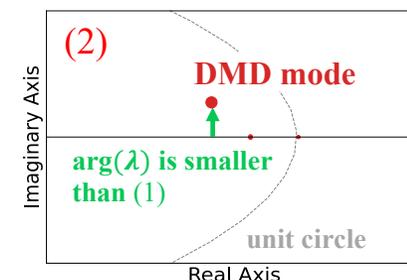
Snapshots



$t_c = 4031$



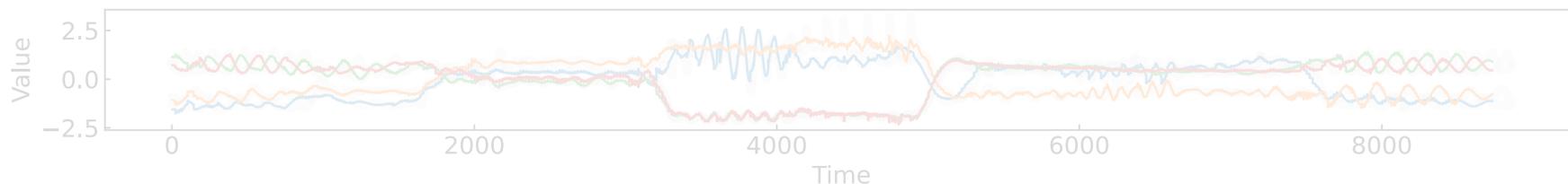
$t_c = 8531$



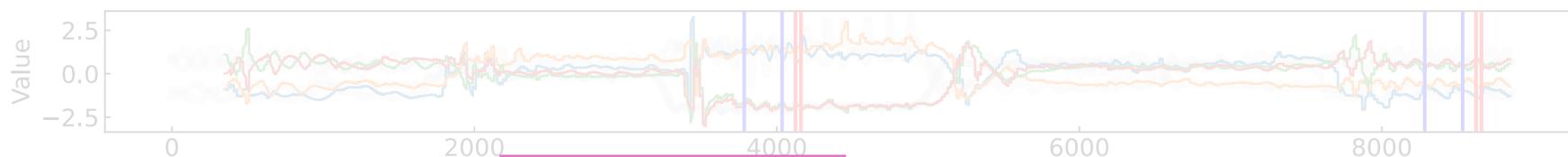
研究背景: 出力例

➤ センサデータストリームに対するリアルタイム予測

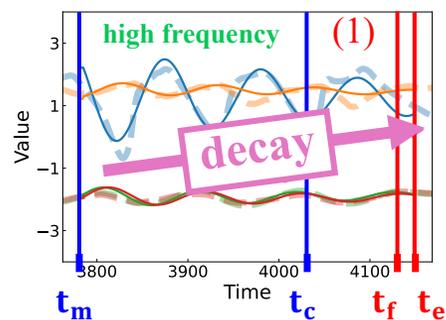
Fitting



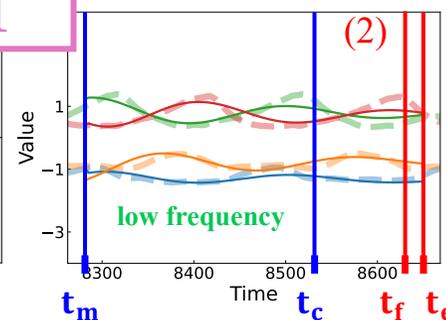
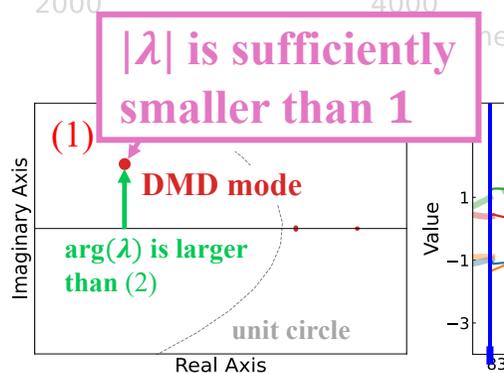
Forecast



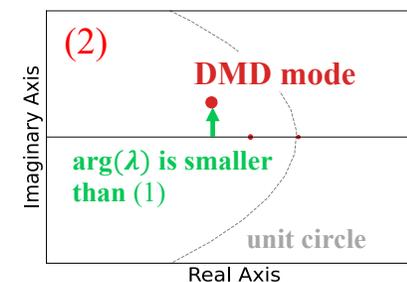
Snapshots



$t_c = 4031$



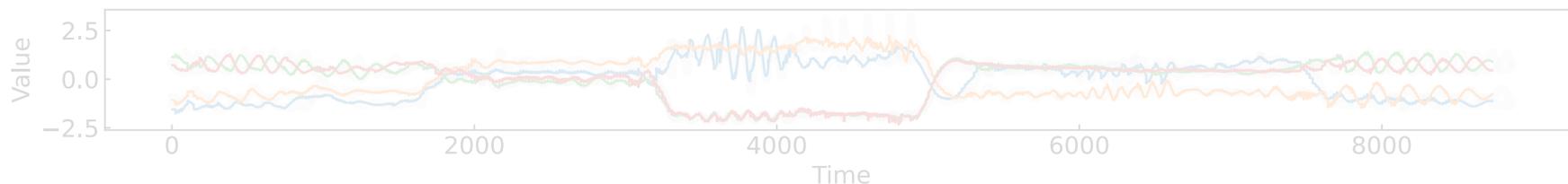
$t_c = 8531$



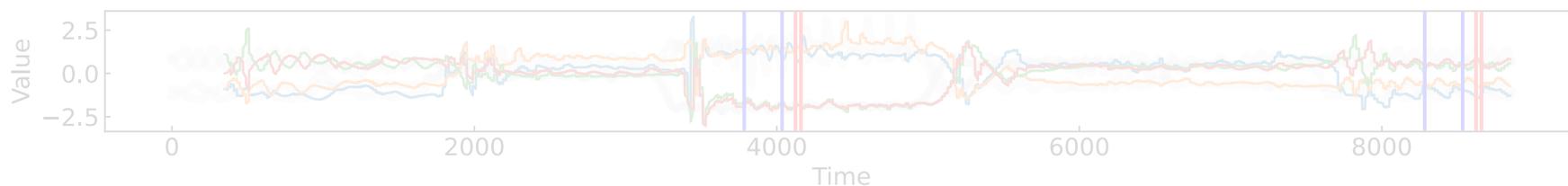
研究背景: 出力例

➤ センサデータストリームに対するリアルタイム予測

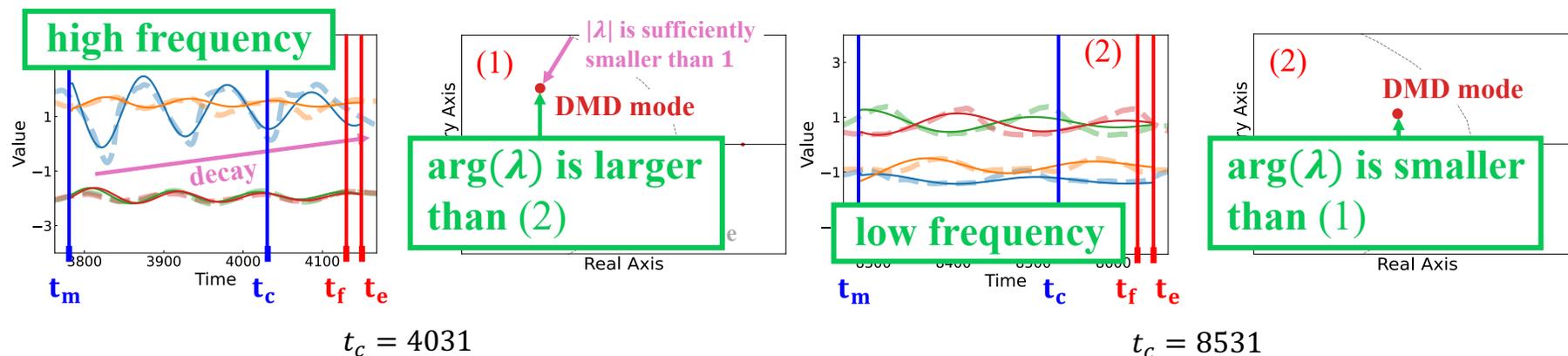
Fitting



Forecast



Snapshots



目次

- ✓ はじめに
- **モデル**
- アルゴリズム
- 実験

モデル

- 達成すべき目標は以下の通り
 - ❖ 時系列データ中の複数のパターン (i.e., レジーム) を捉える
 - ❖ データストリーム X の要約を数式的に表現する

モデル

- 達成すべき目標は以下の通り
 - ❖ 時系列データ中の複数のパターン (i.e., レジーム) を捉える
 - ❖ データストリーム X の要約を数式的に表現する



1. 観測データの**遅延座標**への射影
2. 潜在的な**時間的ダイナミクス**
3. レジームの**動的遷移**

モデル: 遅延座標への射影

➤ 高精度なダイナミクスの抽出のために次元を拡張する

❖ 行列 $\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ から以下を得る

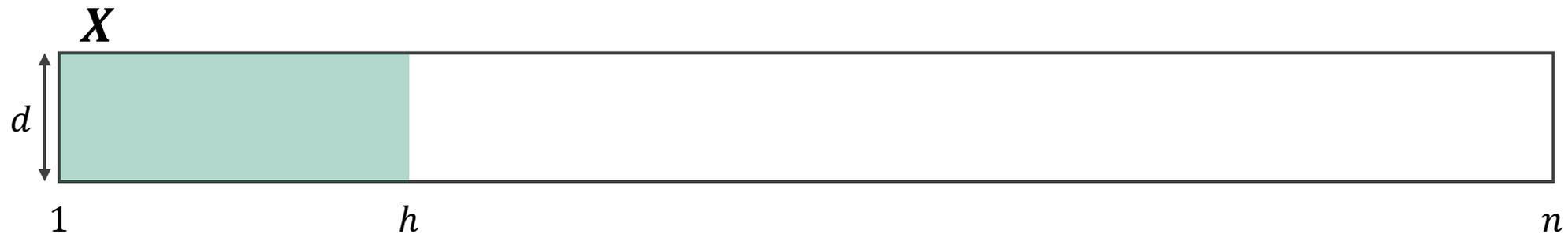
$$\mathbf{X}_{\text{aug}} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_{n-h+1} \\ \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_{n-h+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_h & \mathbf{x}_{h+1} & \cdots & \mathbf{x}_n \end{bmatrix}$$

モデル: 遅延座標への射影

➤ 高精度なダイナミクスの抽出のために次元を拡張する

❖ 行列 $X = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ から以下を得る

$$X_{\text{aug}} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_{n-h+1} \\ \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_{n-h+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_h & \mathbf{x}_{h+1} & \cdots & \mathbf{x}_n \end{bmatrix}$$

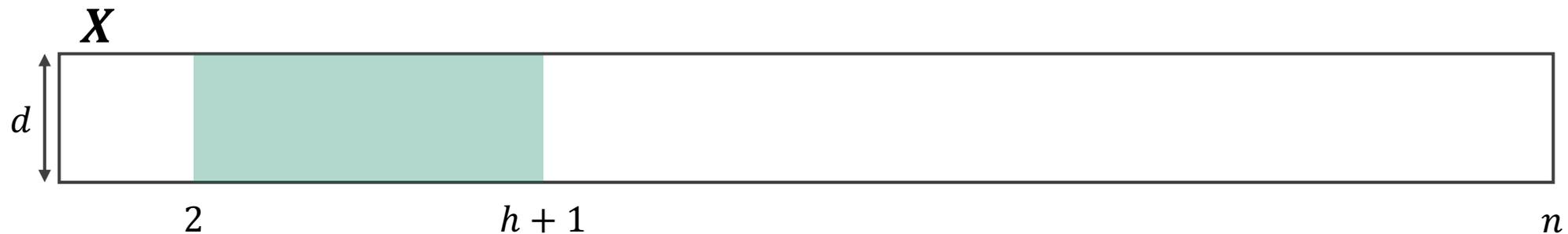


モデル: 遅延座標への射影

➤ 高精度なダイナミクスの抽出のために次元を拡張する

❖ 行列 $\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ から以下を得る

$$\mathbf{X}_{\text{aug}} = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_{n-h+1} \\ \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_{n-h+2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_h & \mathbf{x}_{h+1} & \cdots & \mathbf{x}_n \end{bmatrix}$$



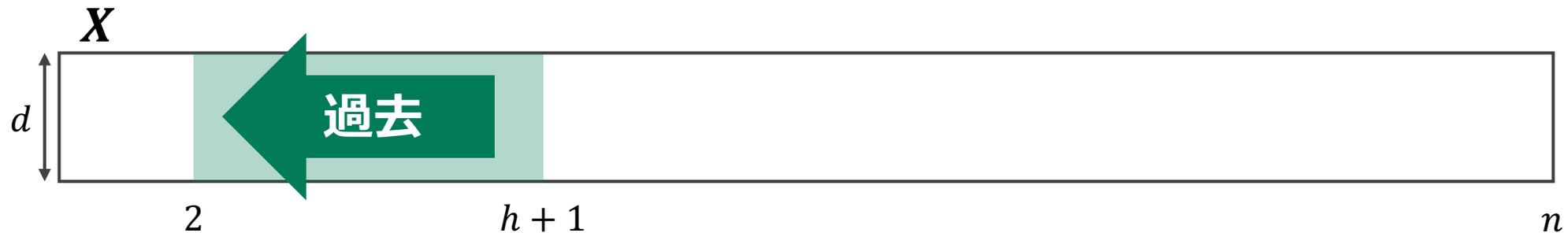
モデル: 遅延座標への射影

➤ 高精度なダイナミクスの抽出のために次元を拡張する

❖ 行列 $X = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ から以下を得る

$$X_{\text{aug}} = \begin{bmatrix} \mathbf{x}_1 & \begin{array}{c} \uparrow \\ \text{過去} \\ \downarrow \\ \mathbf{x}_{h+1} \end{array} & \cdots & \mathbf{x}_{n-h+1} \\ \mathbf{x}_2 & & \cdots & \mathbf{x}_{n-h+2} \\ \vdots & & \ddots & \vdots \\ \mathbf{x}_h & & \cdots & \mathbf{x}_n \end{bmatrix}$$

過去データにより各状態に位相情報が付与される



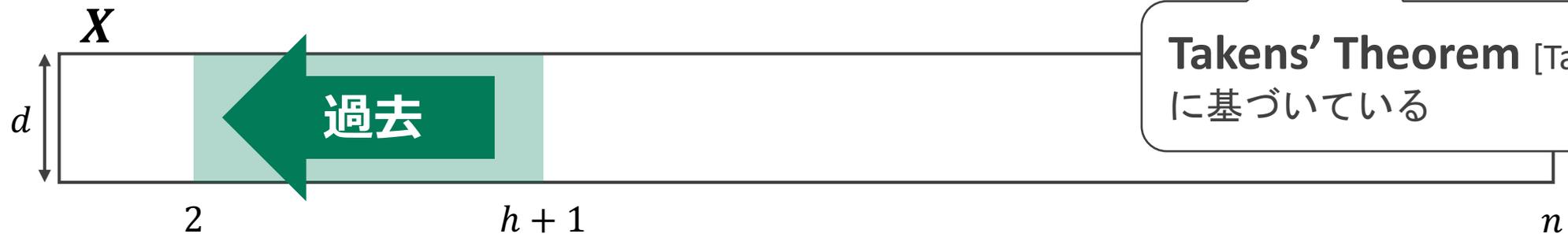
モデル: 遅延座標への射影

➤ 高精度なダイナミクスの抽出のために次元を拡張する

❖ 行列 $X = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_n] \in \mathbb{R}^{d \times n}$ から以下を得る

$$X_{\text{aug}} = \begin{bmatrix} \mathbf{x}_1 & \begin{array}{c} \uparrow \\ \text{過去} \\ \uparrow \\ \mathbf{x}_{h+1} \end{array} & \cdots & \mathbf{x}_{n-h+1} \\ \mathbf{x}_2 & & \cdots & \mathbf{x}_{n-h+2} \\ \vdots & & \ddots & \vdots \\ \mathbf{x}_h & & \cdots & \mathbf{x}_n \end{bmatrix}$$

過去データにより各状態に位相情報が付与される



Takens' Theorem [Takens 1981]
に基づいている

モデル: 潜在的な時系列パターン

- 以下の動的システムにてレジームを表現する

$$\frac{d\mathbf{s}(t)}{dt} = \mathbf{\Lambda}\mathbf{s}(t)$$

$$\mathbf{v}(t) = \mathbf{\Phi}\mathbf{s}(t)$$

| | |
|--|------------------|
| $\mathbf{s}(t) \in \mathbb{C}^k$ | k 次元ベクトルの潜在値 |
| $\mathbf{v}(t) \in \mathbb{R}^{dh}$ | dh 次元ベクトルの推定値 |
| $\mathbf{\Lambda} \in \mathbb{C}^{k \times k}$ | 潜在的な時間ダイナミクス |
| $\mathbf{\Phi} \in \mathbb{C}^{dh \times k}$ | k 次元部分空間への射影行列 |

モデル: 潜在的な時系列パターン

- 以下の動的システムにてレジームを表現する

$$\frac{d\mathbf{s}(t)}{dt} = \mathbf{\Lambda} \mathbf{s}(t)$$
$$\mathbf{v}(t) = \mathbf{\Phi} \mathbf{s}(t)$$

固有値行列

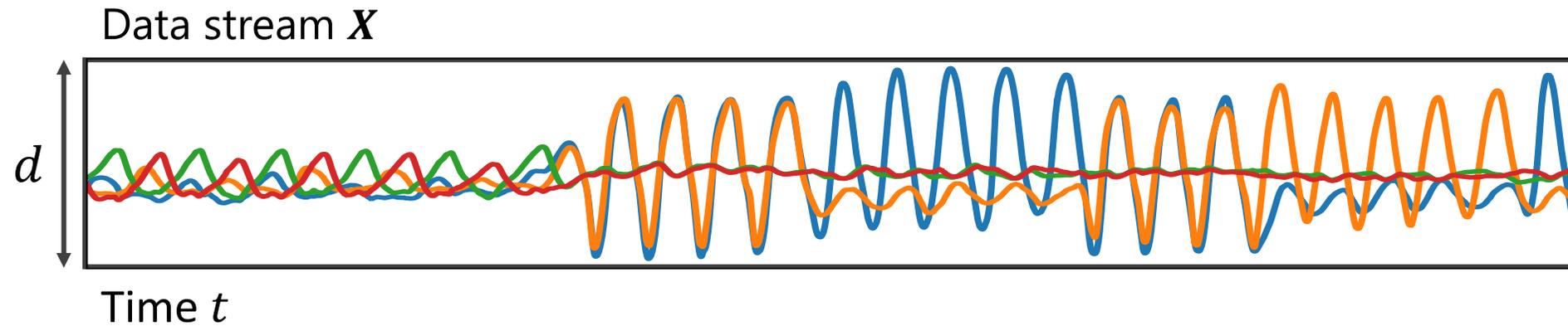
動的モード

| | |
|--|------------------|
| $\mathbf{s}(t) \in \mathbb{C}^k$ | k 次元ベクトルの潜在値 |
| $\mathbf{v}(t) \in \mathbb{R}^{dh}$ | dh 次元ベクトルの推定値 |
| $\mathbf{\Lambda} \in \mathbb{C}^{k \times k}$ | 潜在的な時間ダイナミクス |
| $\mathbf{\Phi} \in \mathbb{C}^{dh \times k}$ | k 次元部分空間への射影行列 |

→ レジーム $\theta = \{\mathbf{\Lambda}, \mathbf{\Phi}\}$

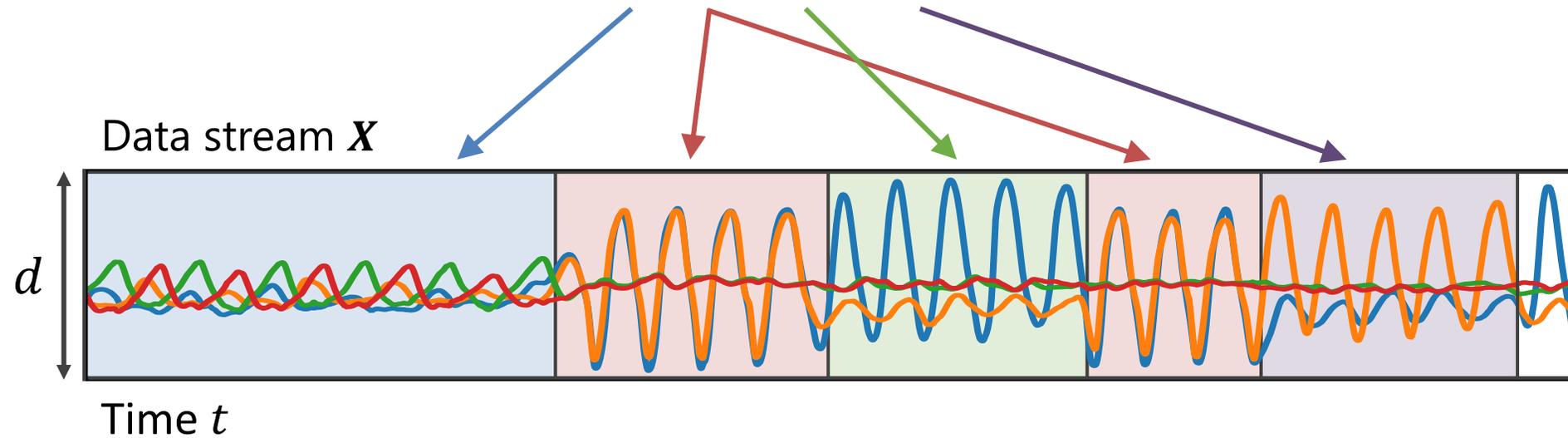
モデル: レジームの動的な遷移

- ▶ ストリーム中の時系列パターンの動的変化
 - ❖ レジームセット $\Theta = \{\theta_1, \theta_2, \dots, \theta_r\}$ ($\theta_i = \{\Phi, \Lambda\}$)



モデル: レジームの動的な遷移

- ▶ ストリーム中の時系列パターンの動的変化
 - ❖ レジームセット $\Theta = \{\theta_1, \theta_2, \dots, \theta_r\}$ ($\theta_i = \{\Phi, \Lambda\}$)

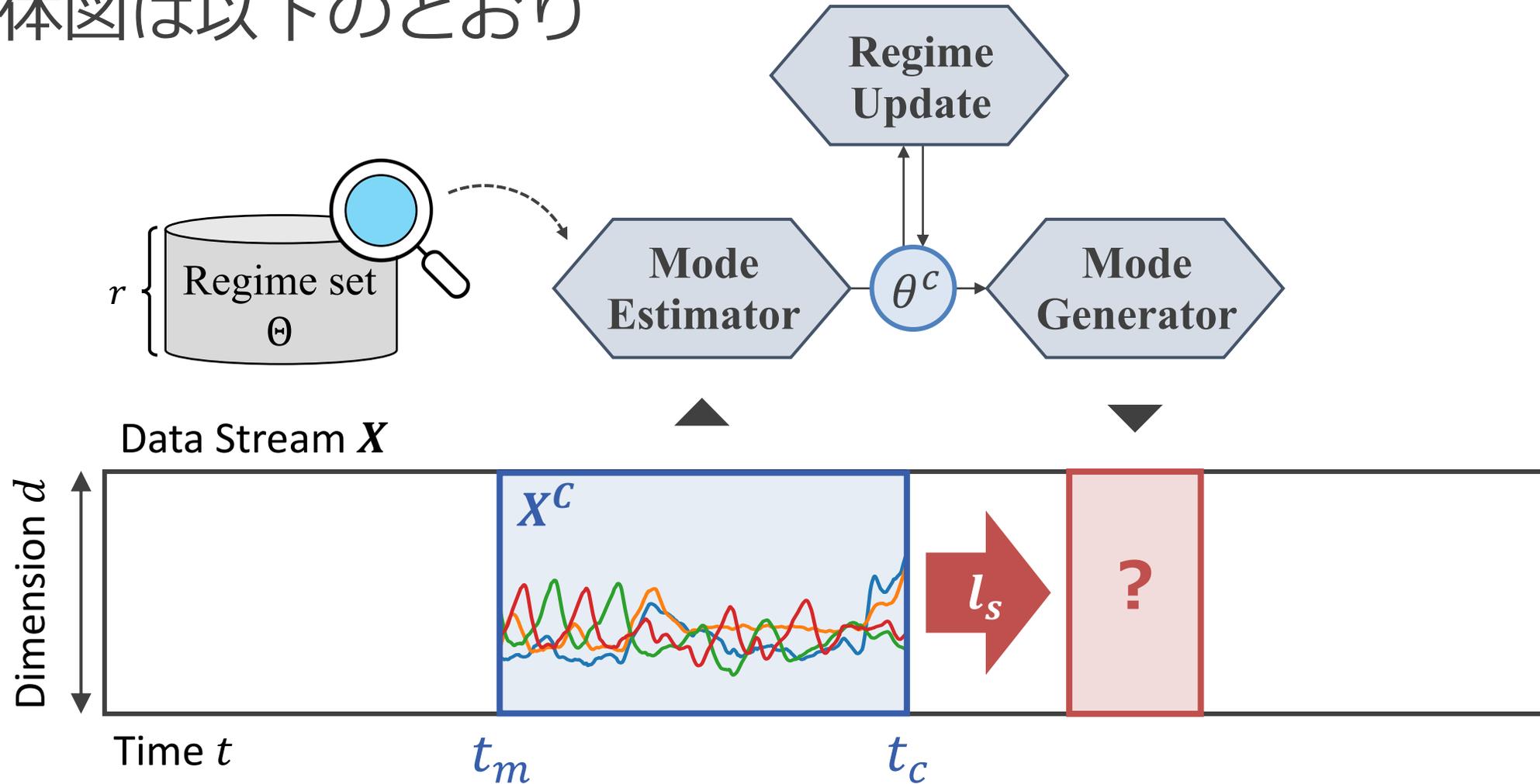


目次

- はじめに
- モデル
- アルゴリズム
- 実験

アルゴリズム

- 全体図は以下のとおり



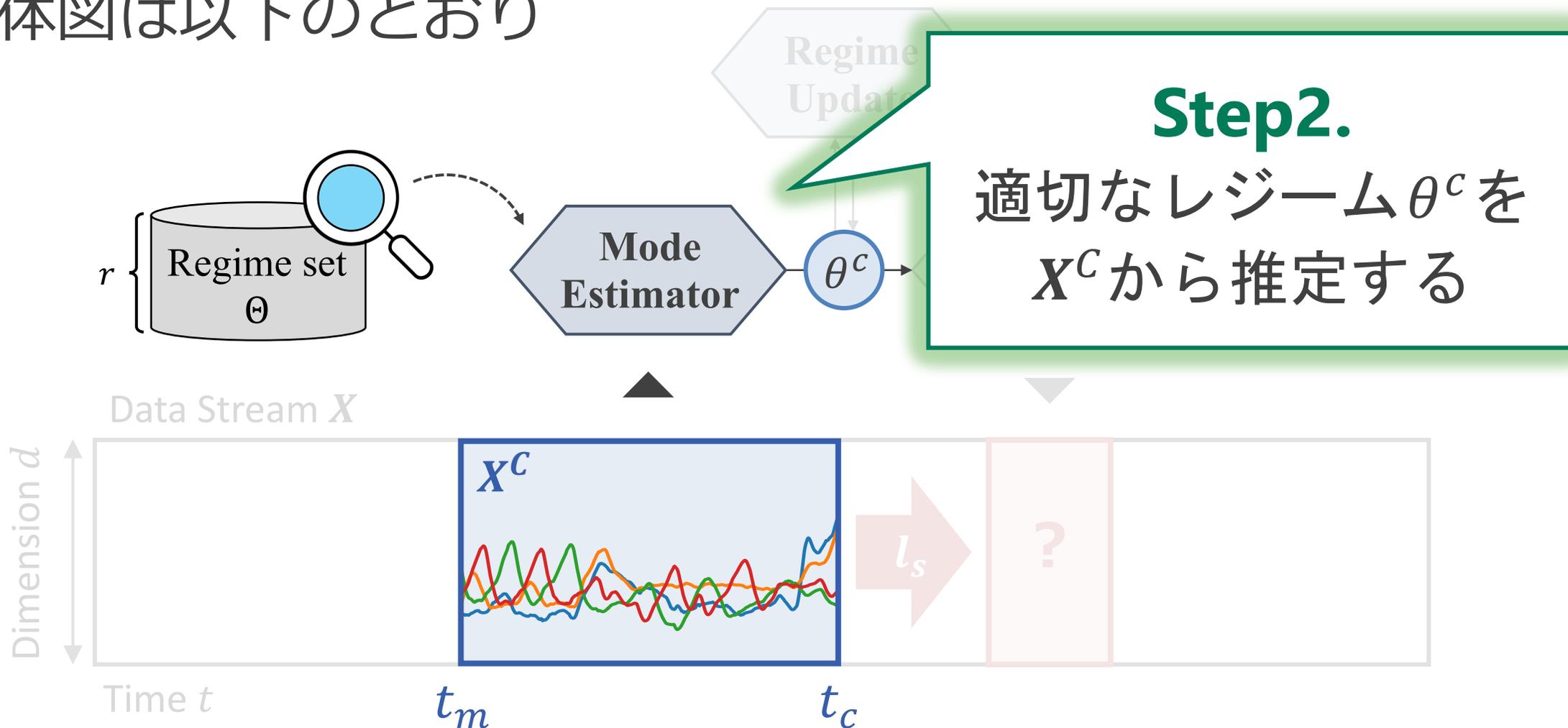
アルゴリズム

- 全体図は以下のとおり



アルゴリズム

- 全体図は以下のとおり



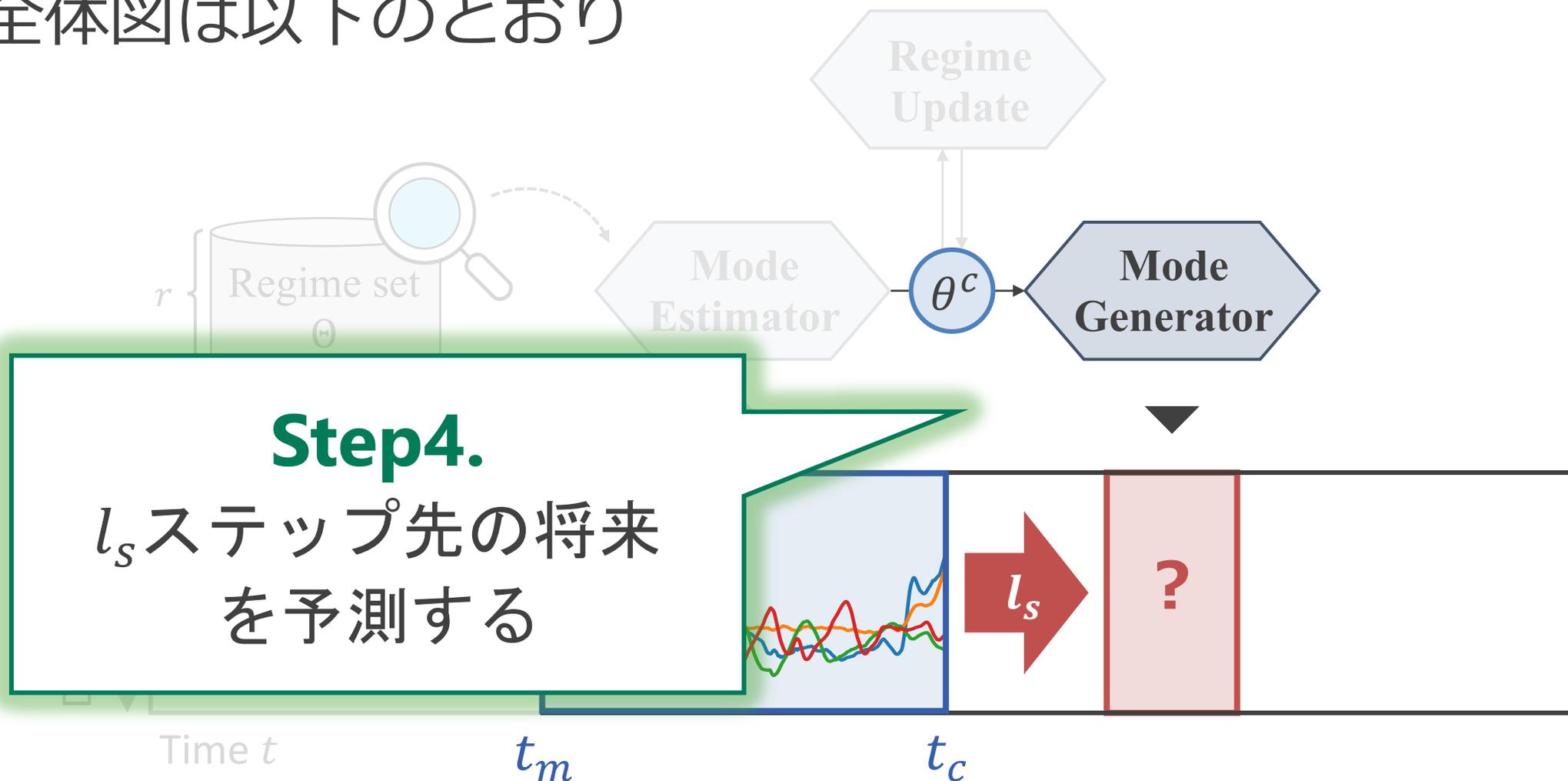
アルゴリズム

- 全体図は以下のとおり



アルゴリズム

- 全体図は以下のとおり



目次

- はじめに
- モデル
- アルゴリズム
- 実験**

実験結果

➤ 以下の疑問に対して回答する

Q1. 正確性

リアルタイム予測に対する提案手法の精度の検証

Q2. 計算時間

データストリームの予測に対する計算時間の検証

実験設定

➤ 9つのデータセット

表2 データセット詳細

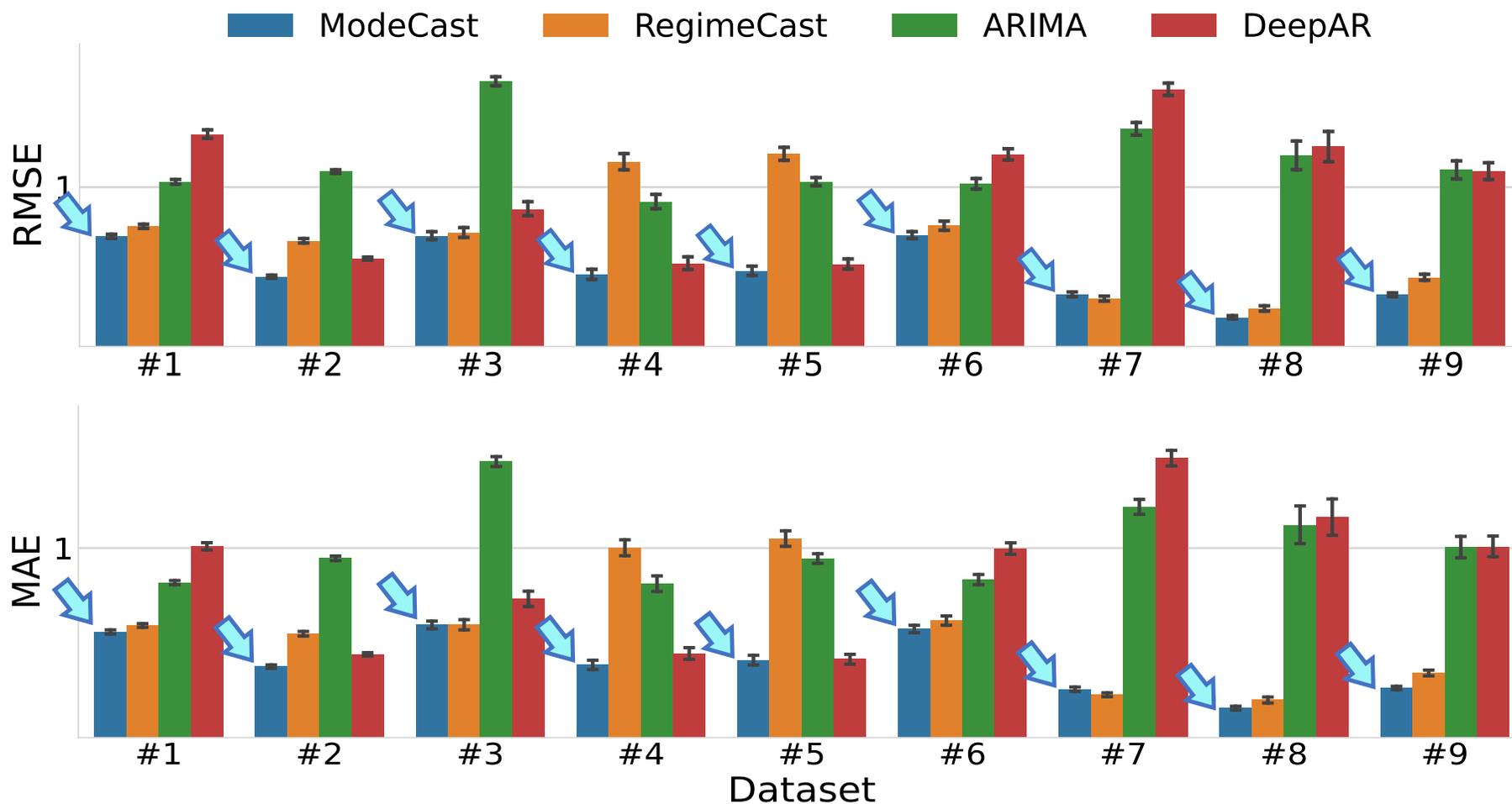
| データセット | 属性・クエリ | 次元 |
|--|--|----|
| センサデータ：CMU Motion Capture Database [37] から取得 | | |
| (#1) <i>exercise</i> | 左右の腕と足の動き | 4 |
| (#2) <i>house-cleaning</i> | ” | 4 |
| (#3) <i>chicken-dance</i> | ” | 4 |
| Web データ：GoogleTrends [38] から取得 | | |
| (#4) <i>G-outdoor</i> | skiing, fishing, cycling | 3 |
| (#5) <i>G-sports</i> | tennis, American football, Major League Baseball (MLB) | 3 |
| (#6) <i>Beer</i> | Corona, Keystone, Coors, Modelo, SierraNevada | 5 |
| (#7) <i>OnlineTV</i> | Netflix, Hulu, YouTube, Amazon(P) | 4 |
| (#8) <i>Social media</i> | Tumblr, Facebook, LinkedIn, SoundCloud, Yelp | 5 |
| (#9) <i>Software</i> | HTML, Java, SQL, JavaScript, HTML5 | 5 |

➤ 3種類の比較手法

- ❖ RegimeCast [Mastubara et al. 2014]
非線形動的システムを活用したリアルタイム予測手法
- ❖ ARIMA [Box and Jenkins 1976]
線形的な関係を元にした古典的な時系列将来予測手法
- ❖ DeepAR [Salinas et al. 2020]
時系列予測のための最新のニューラルネットワーク手法

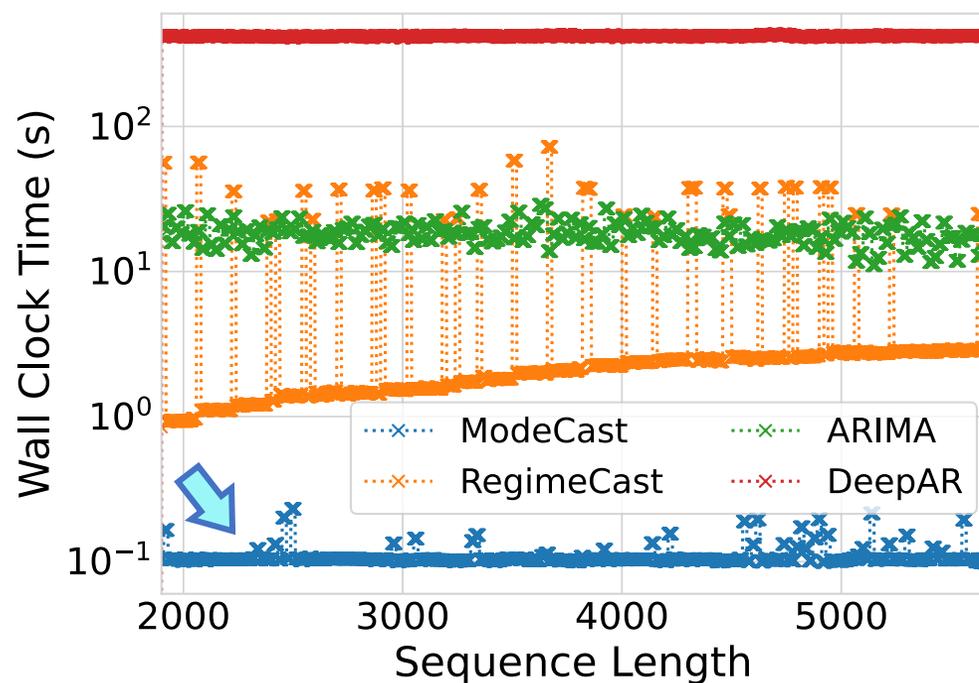
Q1. 正確性

➤ オリジナルデータと予測値の誤差（低いほど良）

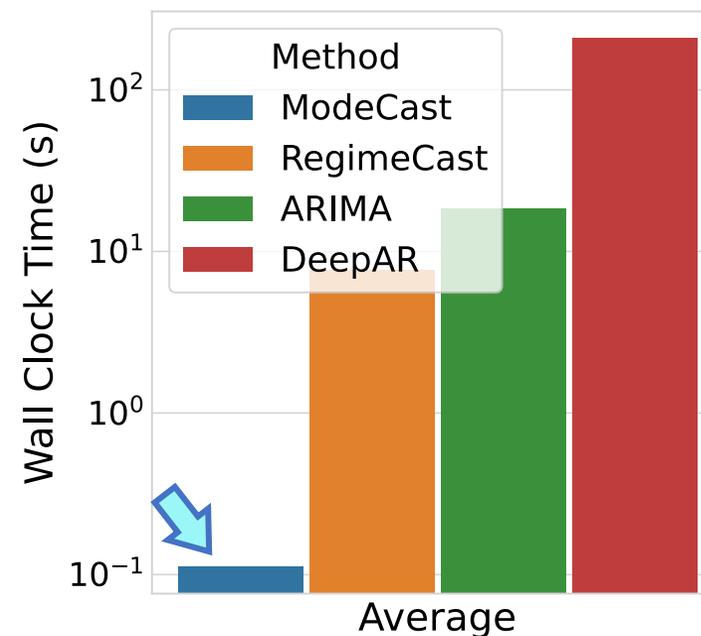


Q2. 計算時間

各時刻における計算時間



平均計算時間



0(1) の時間で予測可能

まとめ

ModeCast は以下の全ての特性を満たす

➤ Effective

- ❖ 固有値に基づいた潜在的な時系列パターンを表現する
- ❖ 複数のレジームを捉え、高精度な将来予測が可能である

➤ General

- ❖ 様々なデータに対して、実用的である

➤ Scalable

- ❖ 計算コストはデータストリームの長さに依存しない
- ❖ 高速に処理が可能である

Appendix

関連研究

- RegimeCast [Mastubara et al. 2014]
 - リアルタイム将来予測を実現
 - 周期性やその減衰率に基づいたパターンの把握が困難 😞
- 深層学習ベースの将来予測手法
 - DeepAR [Salinas et al. 2020] 等
 - モデルの学習に膨大な時間を要するため、最新の情報を考慮したリアルタイムかつ連続的な将来予測が困難 😞

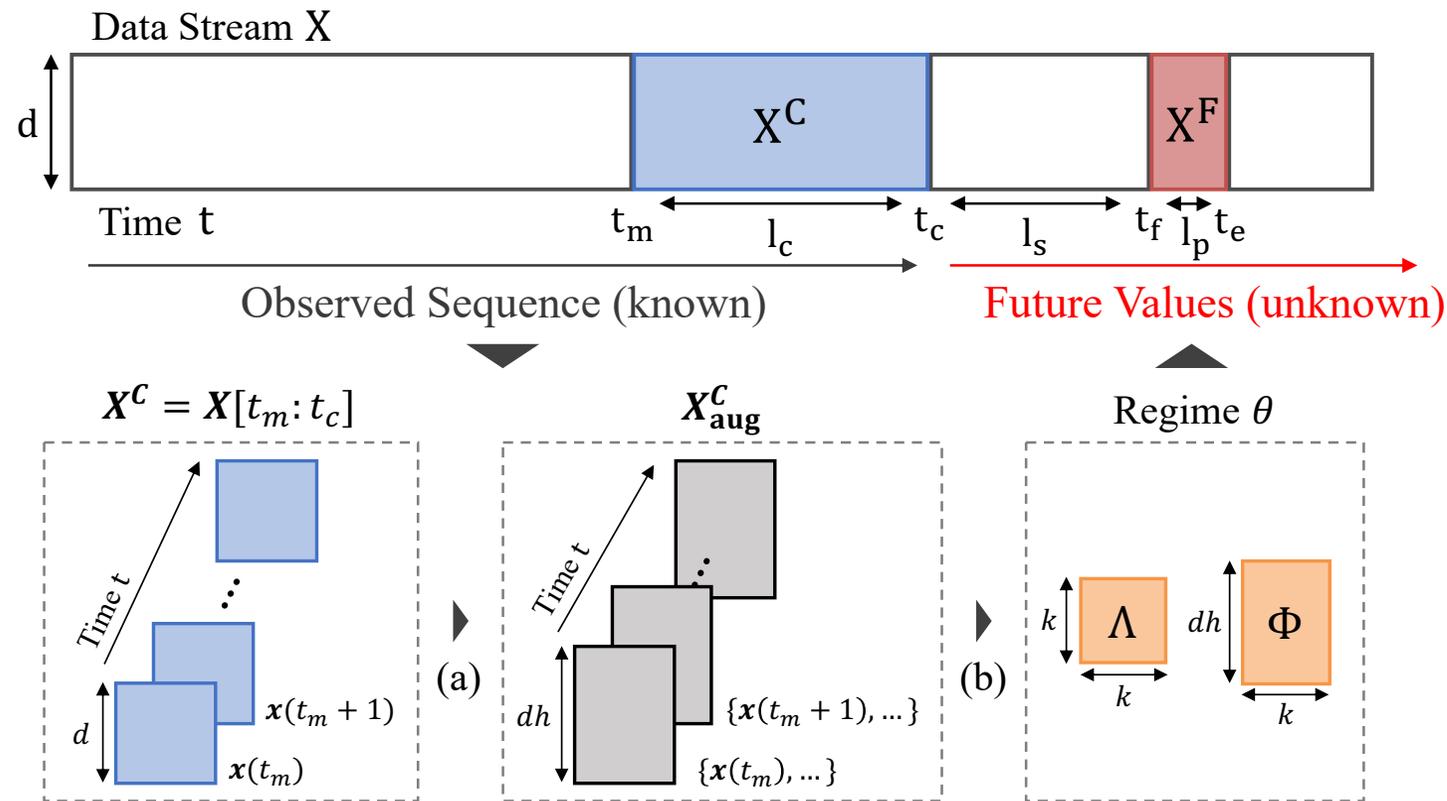
関連研究

➤ 動的モード分解とその応用手法

- 数値流体力学の分野にて初めて提案された [Schmid 2010]
- 物理モデルに関する事前情報が不要
- 多数の分野にて活用されている
 - 動画処理 [Grosek et al. 2014]
 - 脳波解析 [Rohleff 2020]
 - 疫学 [Proctor et al. 2015]
- [Tirunagari et al. 2017] は時系列予測のための手法であるが、単変量データにしか対応していない 😞

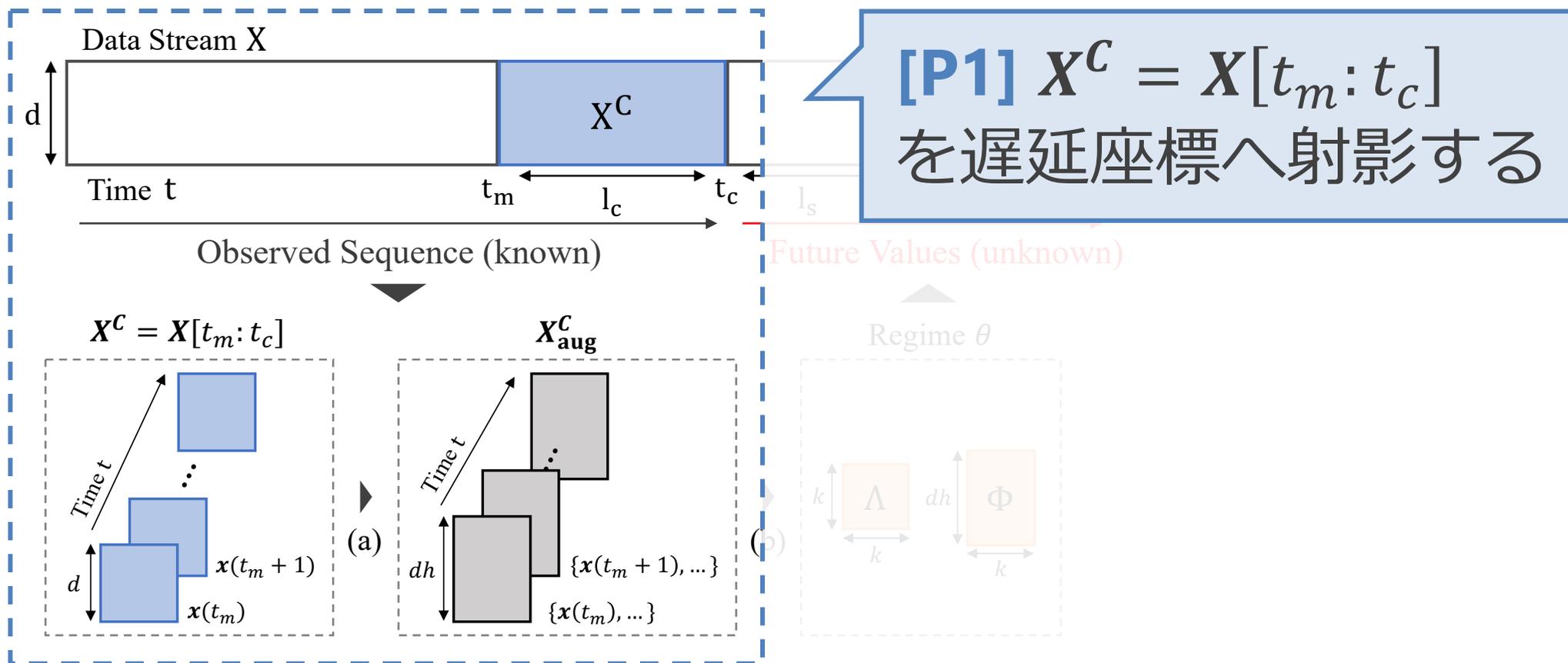
モデル: レジームについて

➤ レジームとは？



モデル: レジームについて

➤ レジームとは？



モデル: レジームについて

➤ レジームとは？

[P2] X_{aug}^C を用いて
レジーム $\theta = \{\Lambda, \Phi\}$
を推定する

