多重マルコフ連鎖を用いた経時確率密度の推定

株式会社タクミインフォメーションテクノロジー
斎藤和宏

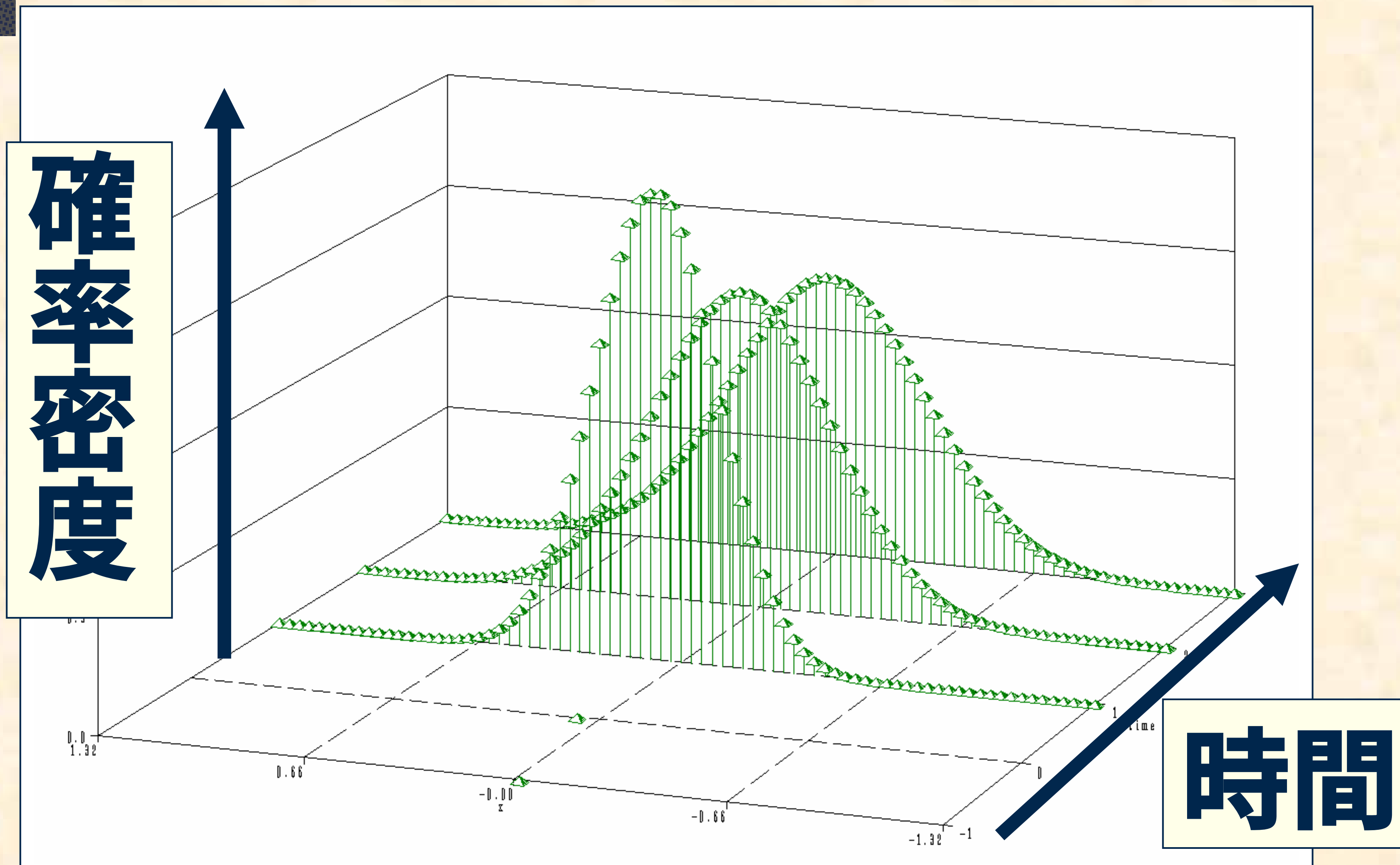
要旨

- 過去のm期の影響を受け、n期先まで推移していく確率密度の推定。
- SAS/IMLを用いて確率分布が次第に拡散していく過程を表示するプログラムを作成した。

目的

- N期先までの特性値ではなく、確率密度の推定。
- 予測する場合に、それよりも前の情報を、できるだけ多く無く使う。

推定された確率密度



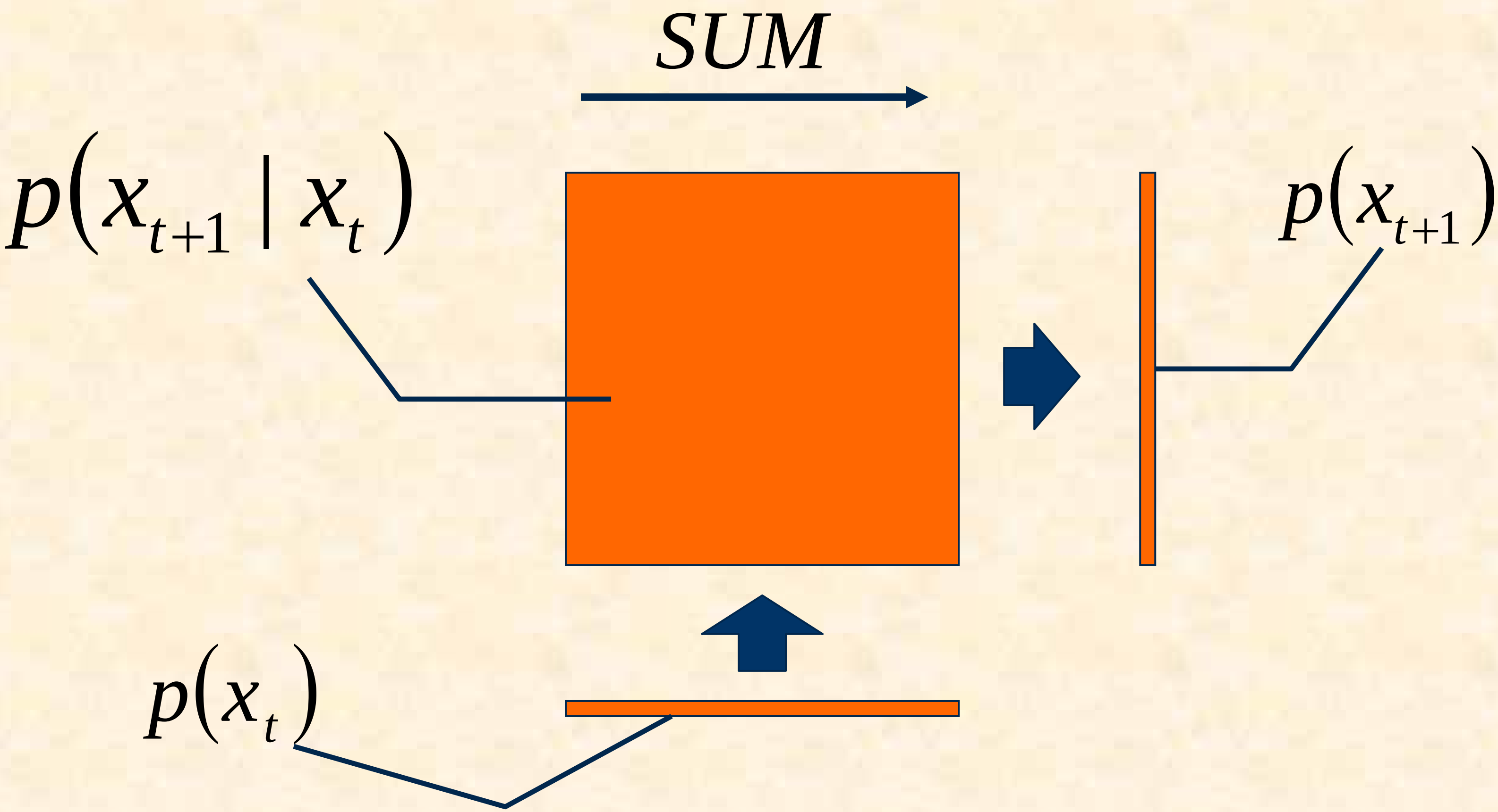
推定の説明に用いる理論

- 多重マルコフ連鎖（M重マルコフ連鎖）
- 説明変数の確率密度を考慮した回帰分析

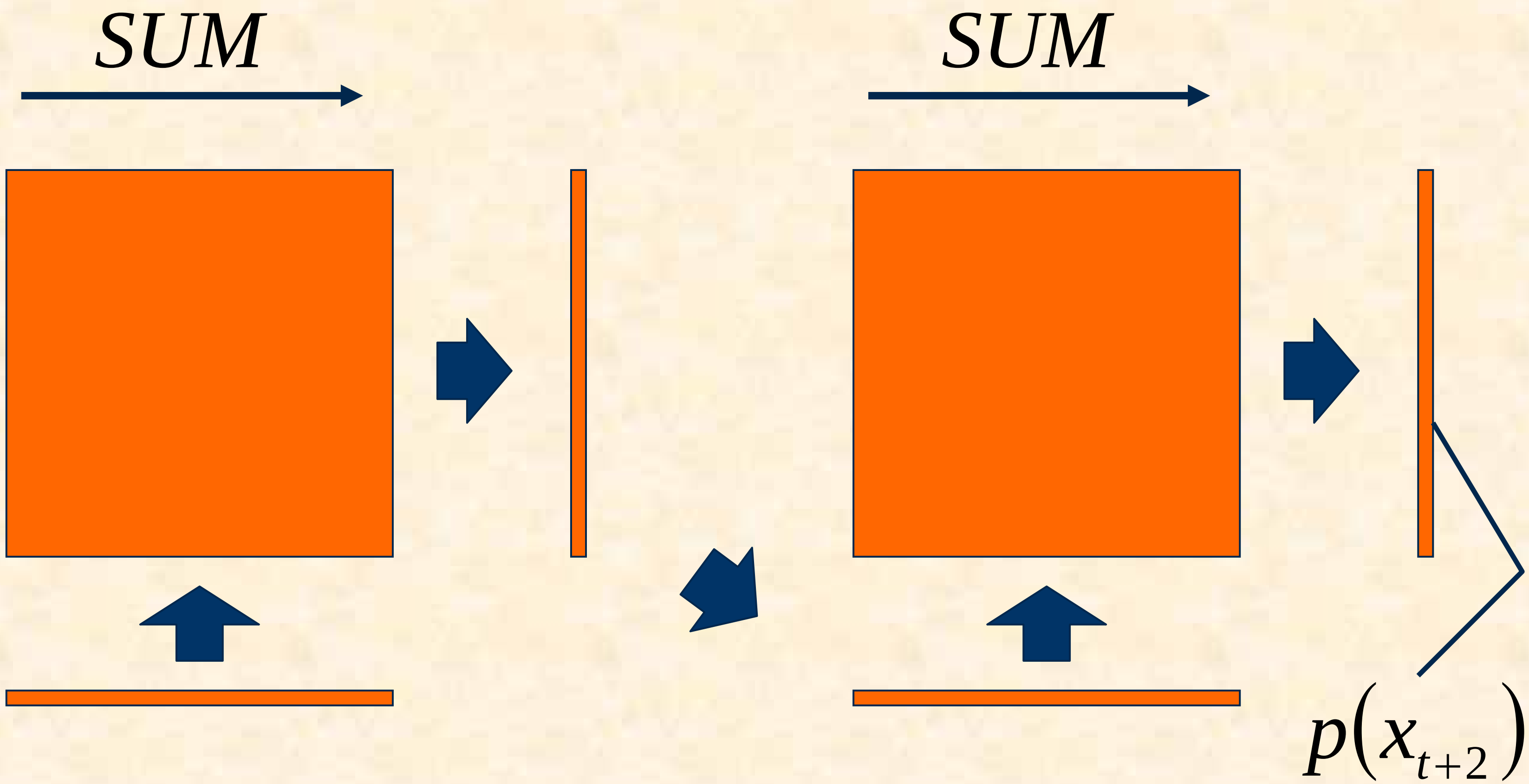
多重マルコフ連鎖

- 通常のマルコフ連鎖は一つ前の状態を受けている。
- 多重マルコフ連鎖では一つではなく複数の期間の影響を受けている。

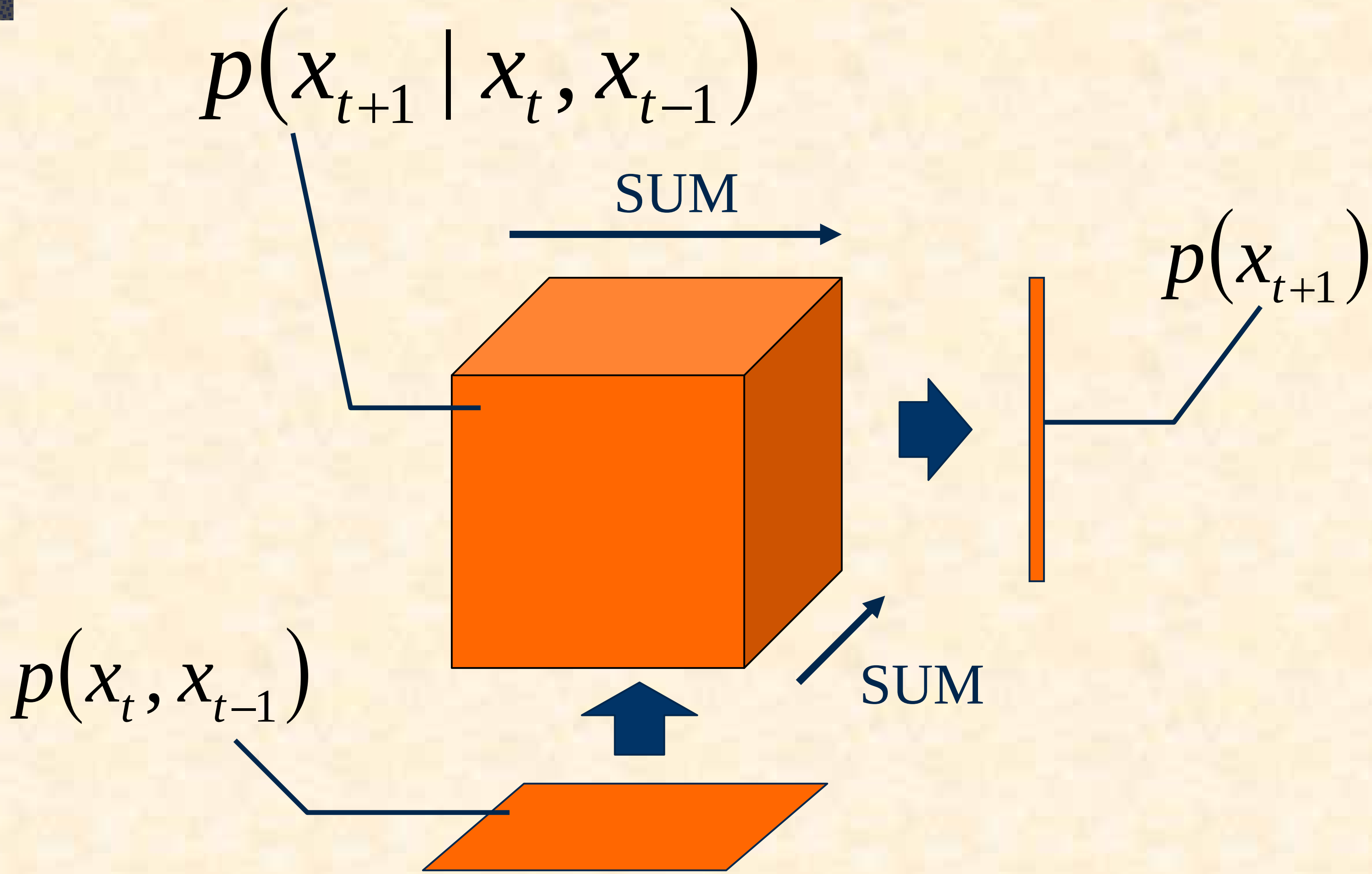
マルコフ連鎖



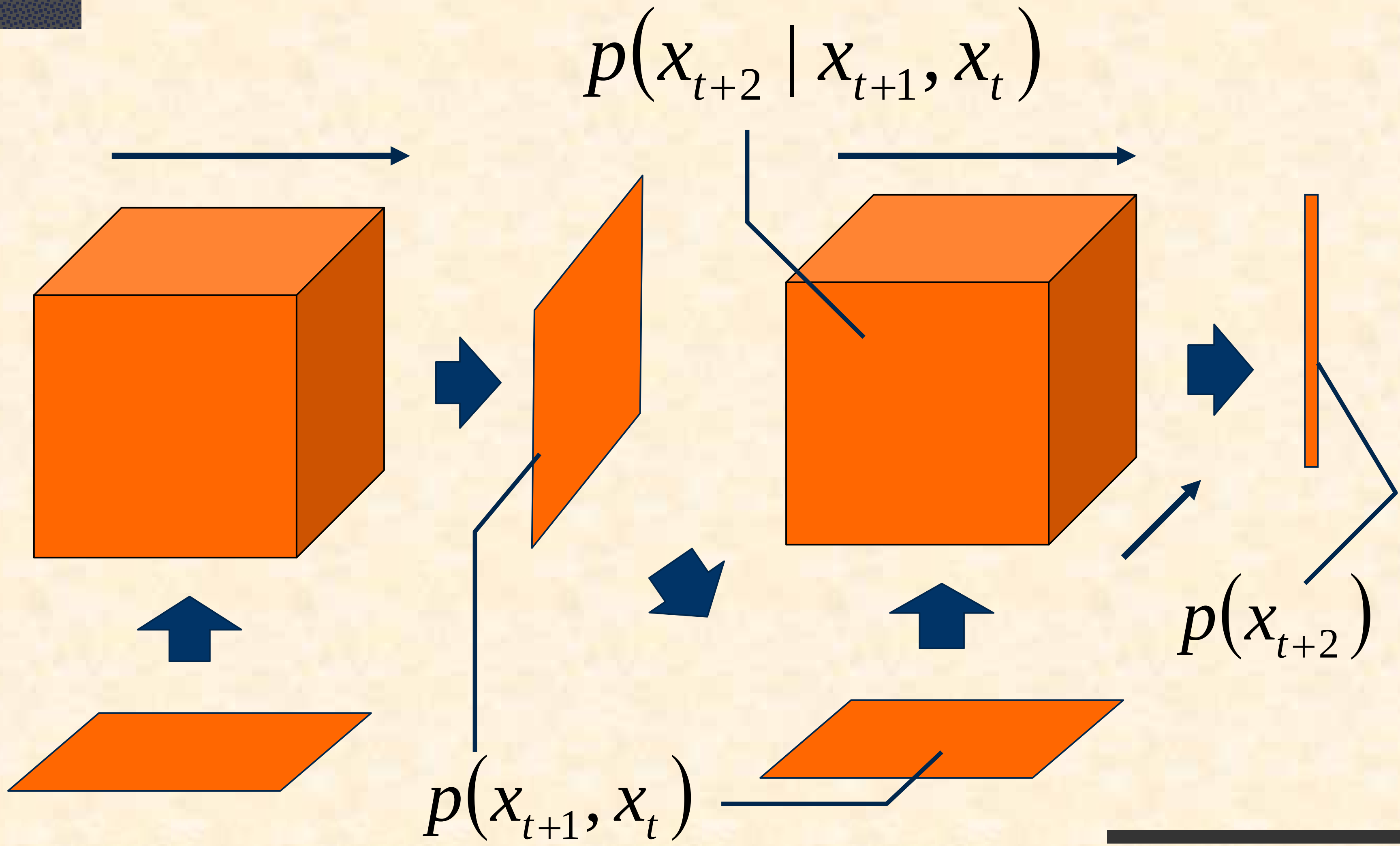
マルコフ連鎖



2重マルコフ連鎖



2重マルコフ連鎖による2期先の推定



2重マルコフ連鎖による2期先の推定

$$p\left(x_{t+2,l}\right)=$$

$$\sum_{k}\sum_{j}\sum_{i} p\left(x_{t+2,l} \mid x_{t+1,k}, x_{t,j}\right)$$

$$p\left(x_{t+1,k} \mid x_{t,j}, x_{t-1,i}\right)p^{*}\left(x_{t,j}, x_{t-1,i}\right)$$

予測の対象としての確率密度

- ある説明変数が与えられた時の目的変数の確率密度を予測の対象とする。
- 確率密度は確率としての意味を成していないが、dyを掛け合わせることで、微小区間での確率となる。

$$\hat{p}(y)$$

確率密度

⟺

$$\hat{p}(y)dy$$

微小区間での確率

説明変数自体の確率密度

- 回帰分析では、ある定まった値を説明変数としているが、実際にはその値自体がわからない場合がある。
- この確定されていない状態を表すためには確率密度関数を使用することが妥当だと思われる。

説明変数自体の確率密度

- 説明変数を条件とした条件付確率を用いると、予測値の確率密度に説明変数の確率密度を反映することができる。

$$\hat{p}(y) = \int_{R^m} p(y \mid \mathbf{x}) \cdot p^*(\mathbf{x}) \mathbf{dx}$$

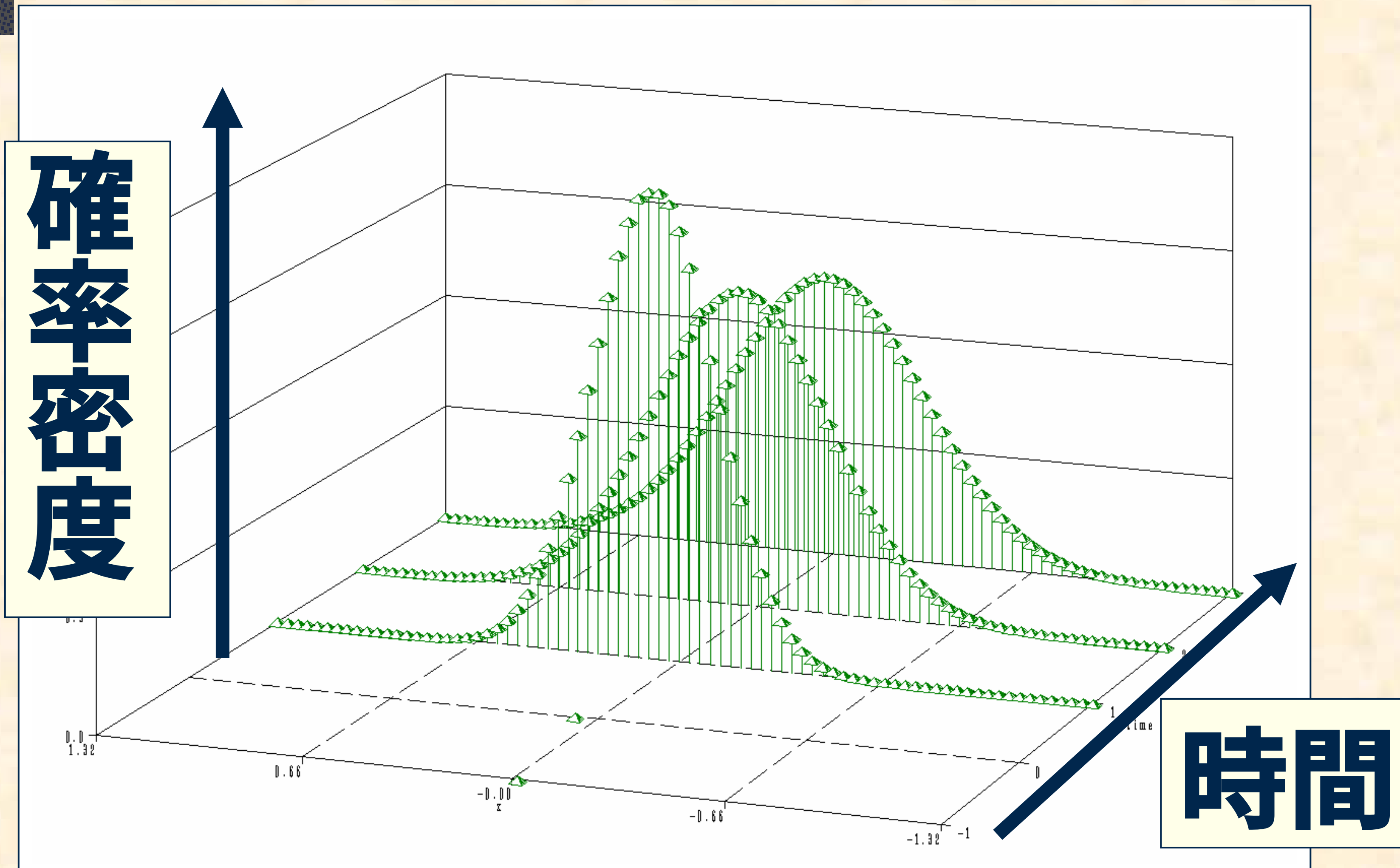
- 予測値はyに関して積分することで求めることができる。

$$\hat{y} = \int y \cdot \hat{p}(y) dy$$

確率密度の推移

- 過去の情報を用いて、後の状態を予測するには、どのような手法が妥当だろうか？
- 推定の起点から未来に向かうにしたがって分からないことが増えると考えるのが自然である。
- 確率分布ならば、経時的に分散が大きくなっていくことで表現することができる。

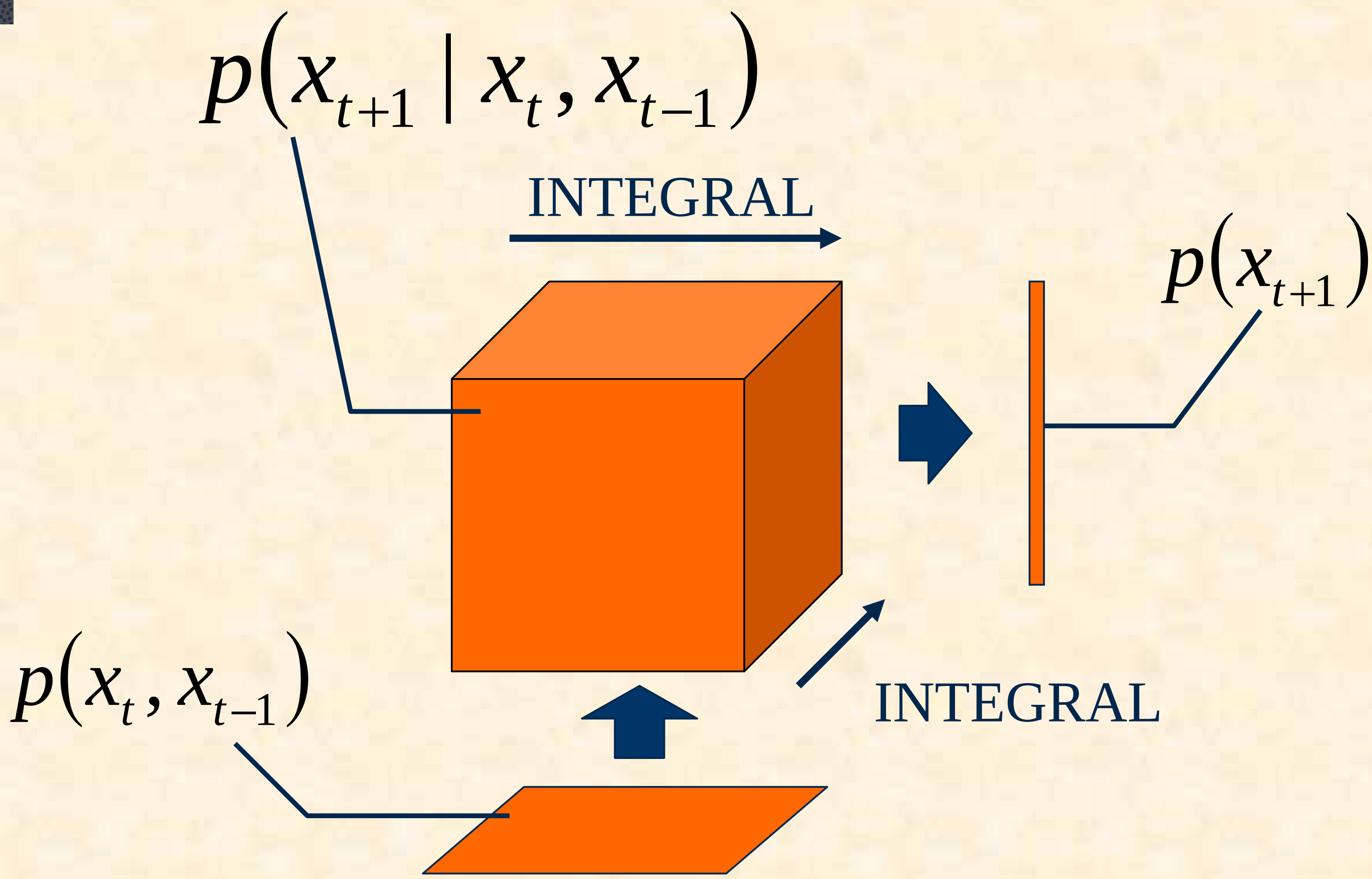
確率密度の推移（3期先まで）



確率密度の推移

- 特に拡散していく過程が正規分布しているしている必要はない。
- 二山でも歪んでいても、尤もらしい場合も考えられる。

確率密度の推移（1期先）

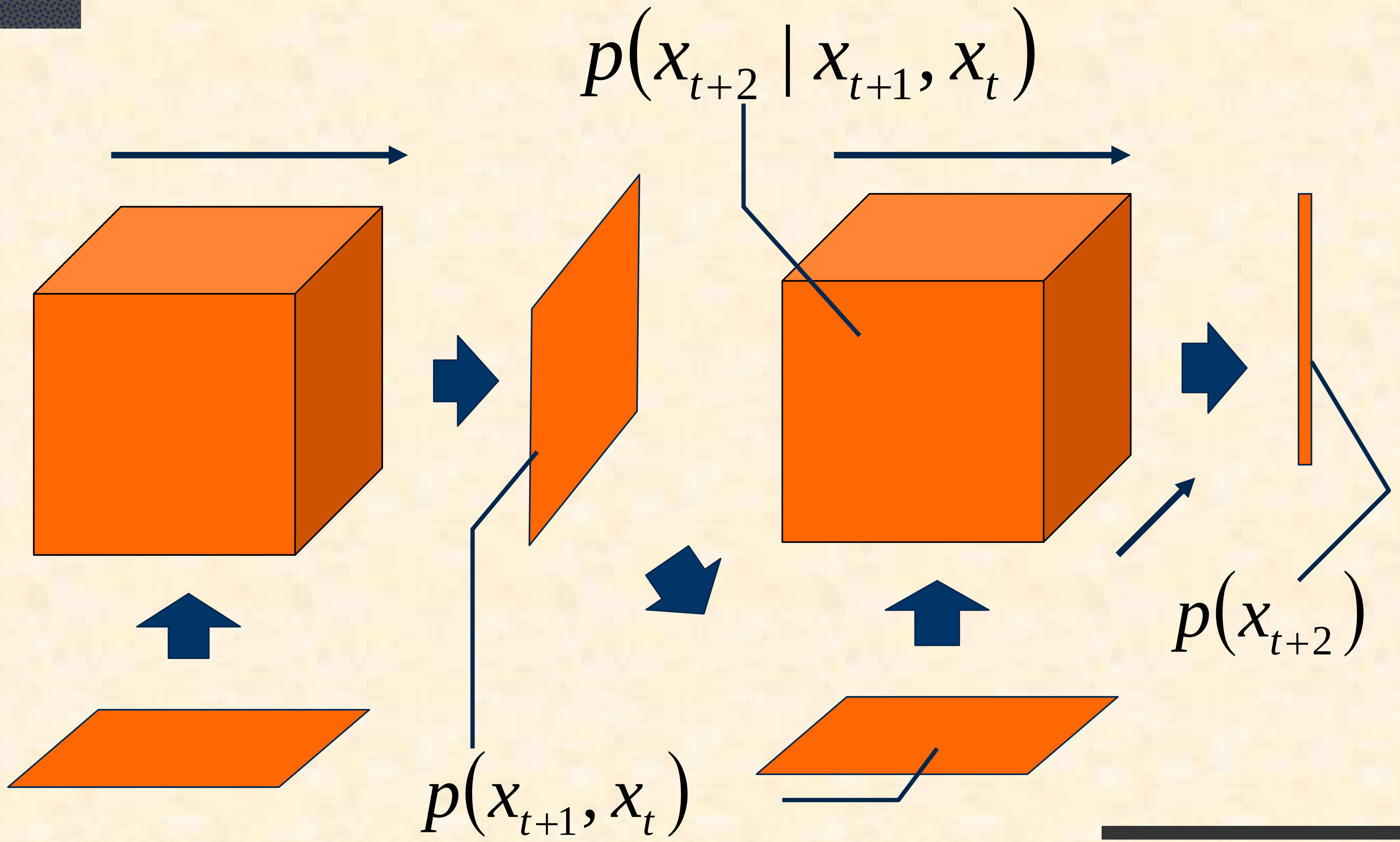

$p\left(x_{t+1} \mid x_t, x_{t-1}\right)$
INTEGRAL
$p\left(x_{t+1}\right)$
$p\left(x_t, x_{t-1}\right)$
INTEGRAL

確率密度の推移（1期先）

$$p_{2,1}(x_{t+1}) =$$

$$\iint \varphi(x_{t+1} \mid x_t, x_{t-1})\, p^*(x_t, x_{t-1})\, dx_{t-1}\, dx_t$$

確率密度の推移（2期先）



確率密度の推移（2期先）

$$p_{2,2}\left(x_{t+2}\right)=$$

$$\iint \underline{\varphi\left(x_{t+2} \mid x_{t+1}, x_{t}\right)} \varphi\left(x_{t+1} \mid x_{t}, x_{t-1}\right) p^{*}\left(x_{t}, x_{t-1}\right) dx_{t-1}\, dx_{t}\, \underline{dx_{t+1}}$$

追加

SAS/IMLによる積分

積分対象のモジュール名

call quad(result, ‘func’, range);

結果が入る変数

積分範囲

-∞～∞　　{.M .P}

G3Dプロシジャによる3Dプロット

- Active Xによるプロット

今後の課題

- 複雑な形の条件付確率や初期状態を用いた分析。そして確率密度が高くなる区間の検出。

- デリバティブの計算ではポピュラーな、Low Discrepancyの一様乱数を用いた重積分。
- これにより、現実的な時間内で重積分を実行することができる。