東京大学 工学部 計数工学科/物理工学科

応用音響学: E3 – DPからHMMへ

嵯峨山 茂樹 <sagayama@hil.t.u-tokyo.ac.jp>

東京大学 工学部 計数工学科 http://hil.t.u-tokyo.ac.jp/

DP マッチング (動的時間軸伸縮 DTW)

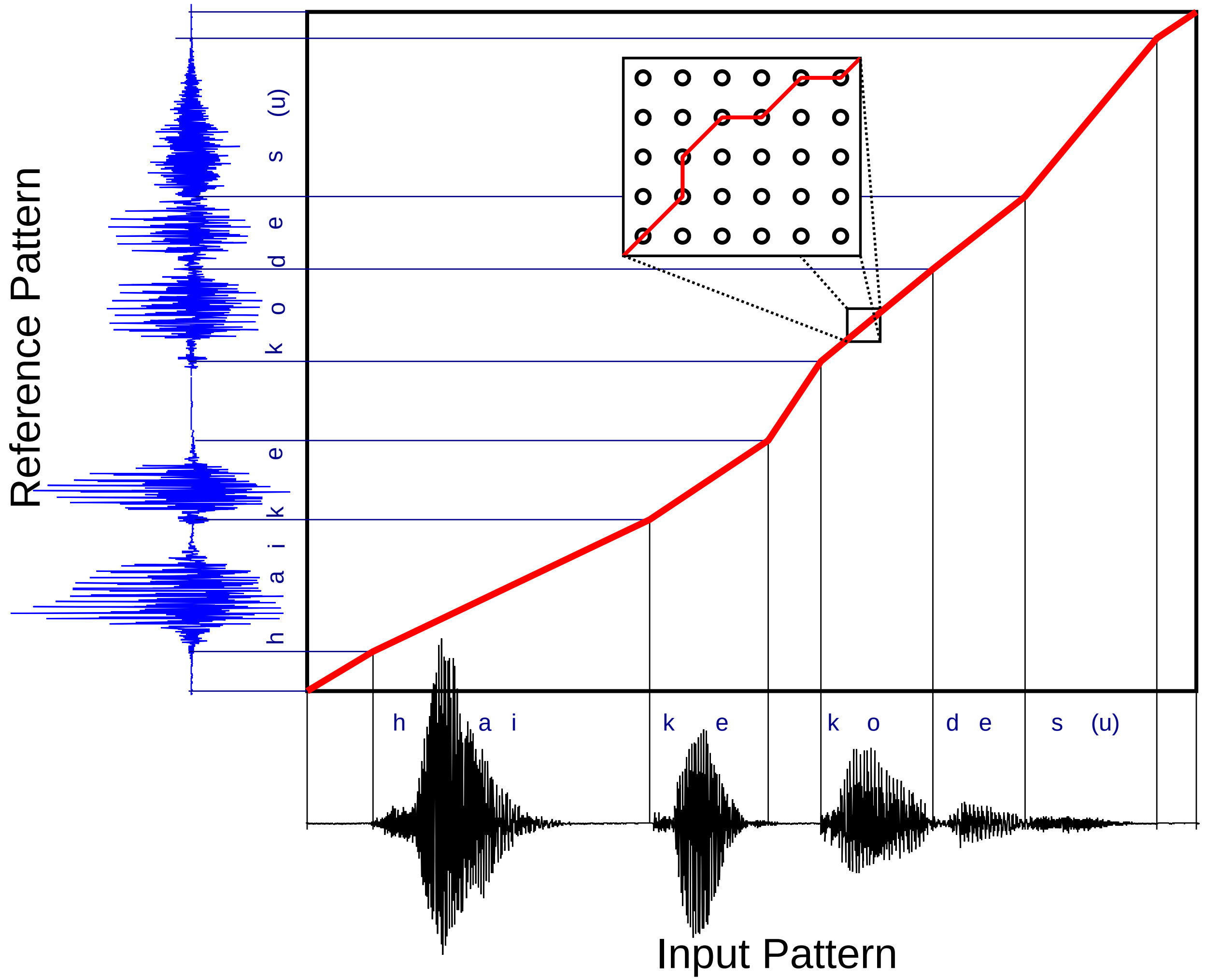

図1. 入力パターンと標準パターンの間の DP matching (Dynamic Time Warping)

節点と枝の両方にコストがある動的計画法の例

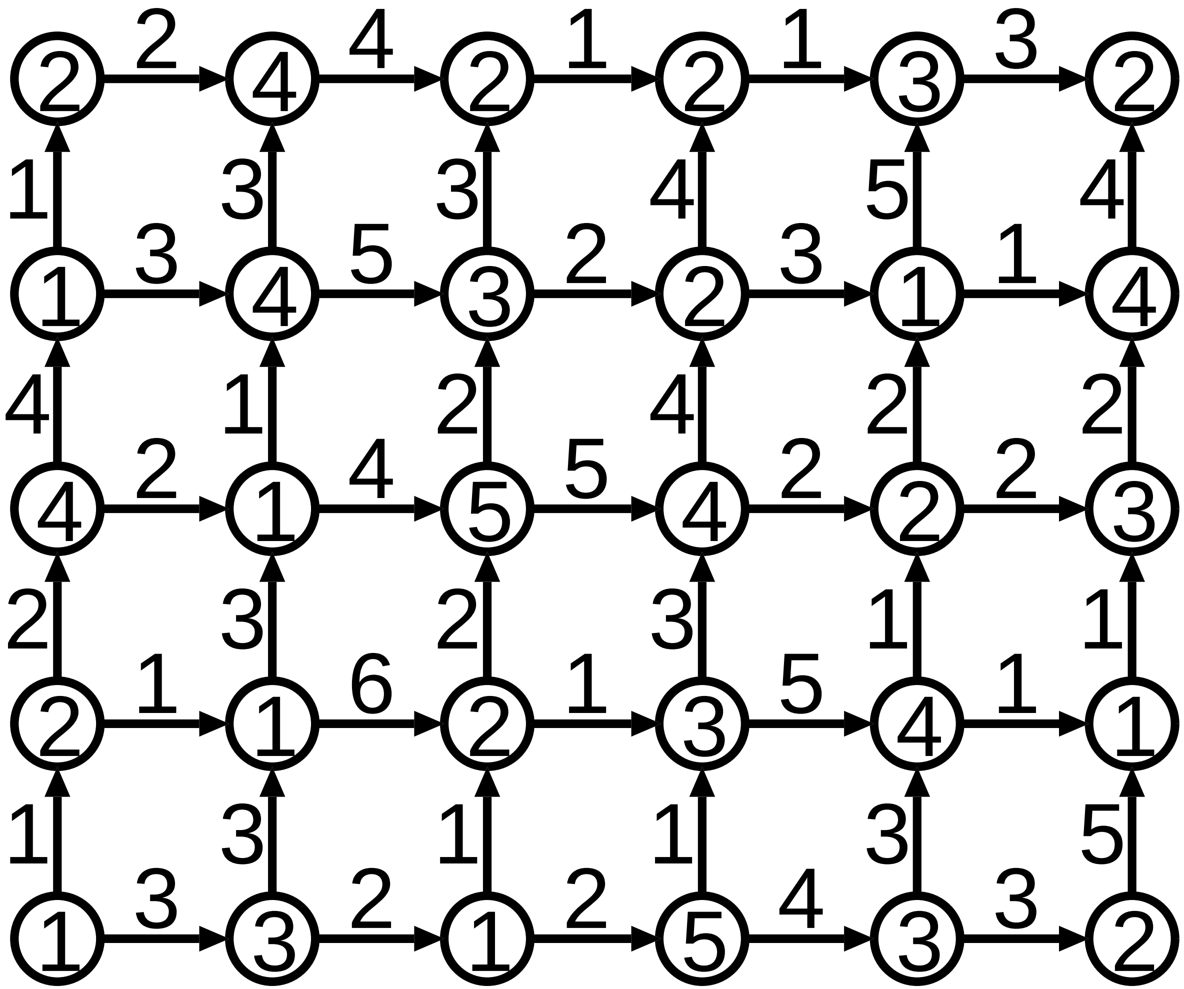

図2. 考え方は同じ：部分最適化問題へ分解できる

節点と枝にコストがある非対称経路DP

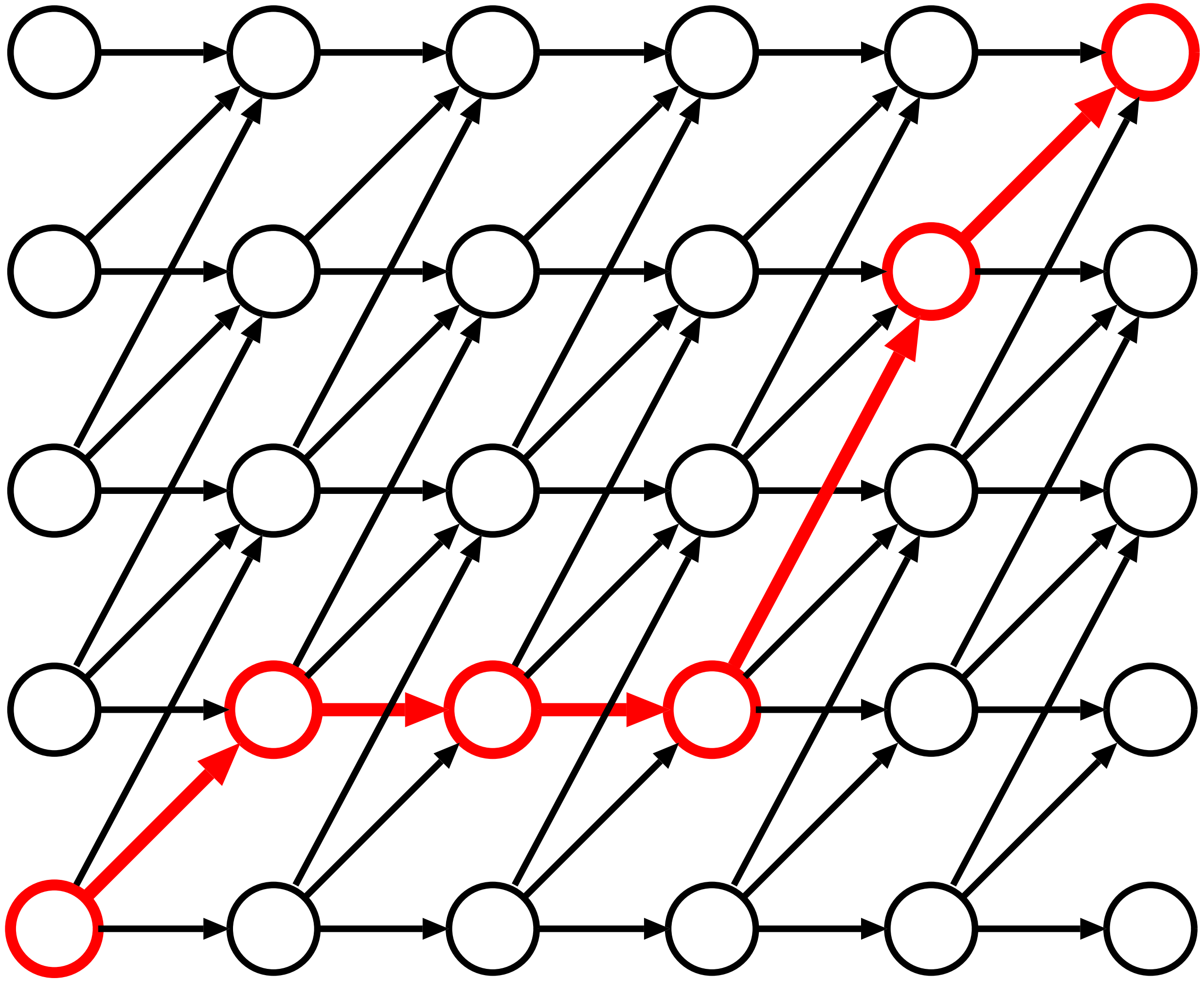

図3. 節点(丸)と枝(矢印)の両方にコスト(距離)が定義されている。

音声特徴ベクトルの軌跡

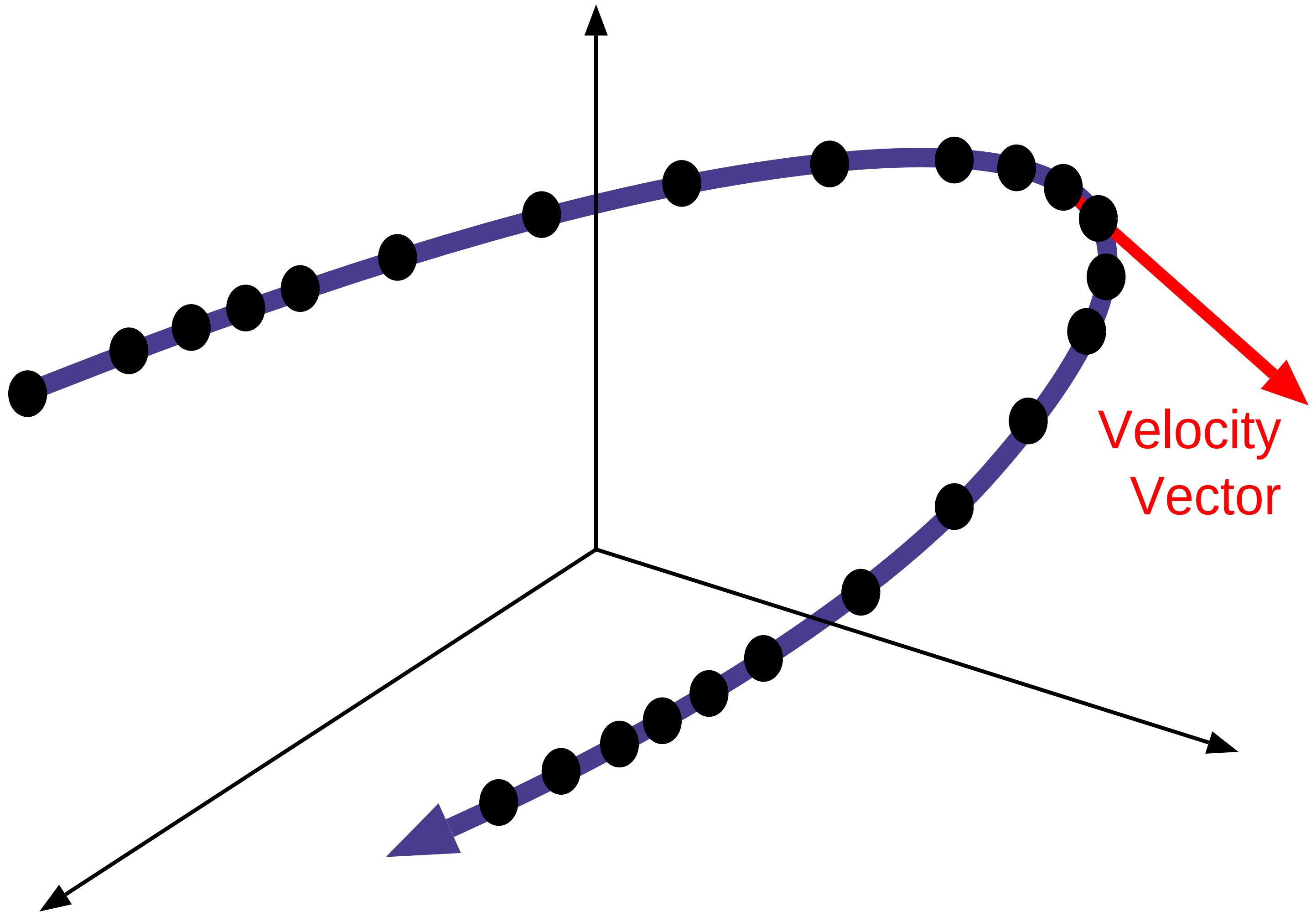

図4. 音声特徴ベクトルの時系列は特徴空間上の点の運動軌跡と捉えることができる。

DPの解釈

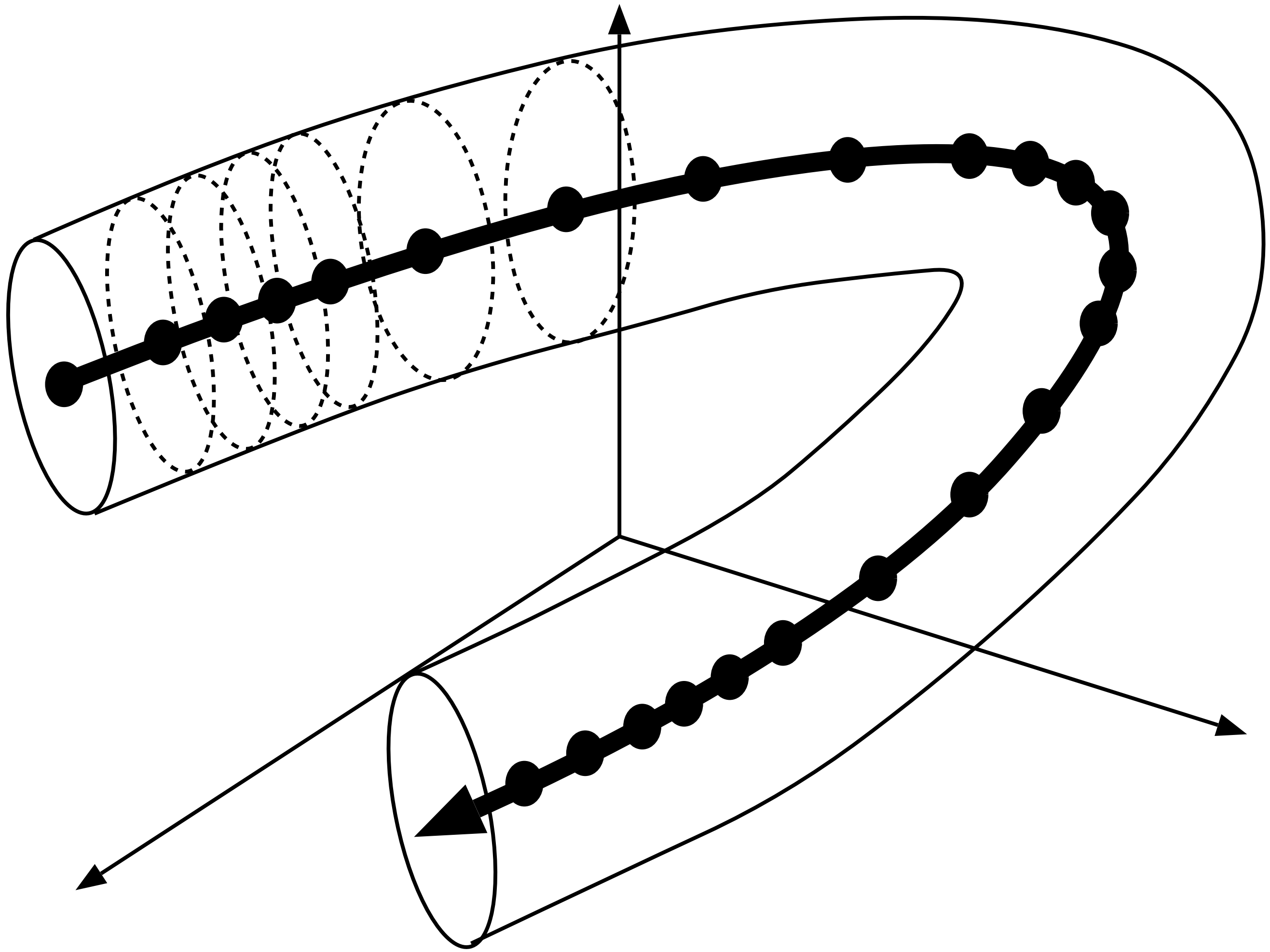

図5. DPは標準パターンを中心線として均一な管を考えて、それによって距離を計算することに相当する。

HMMの解釈

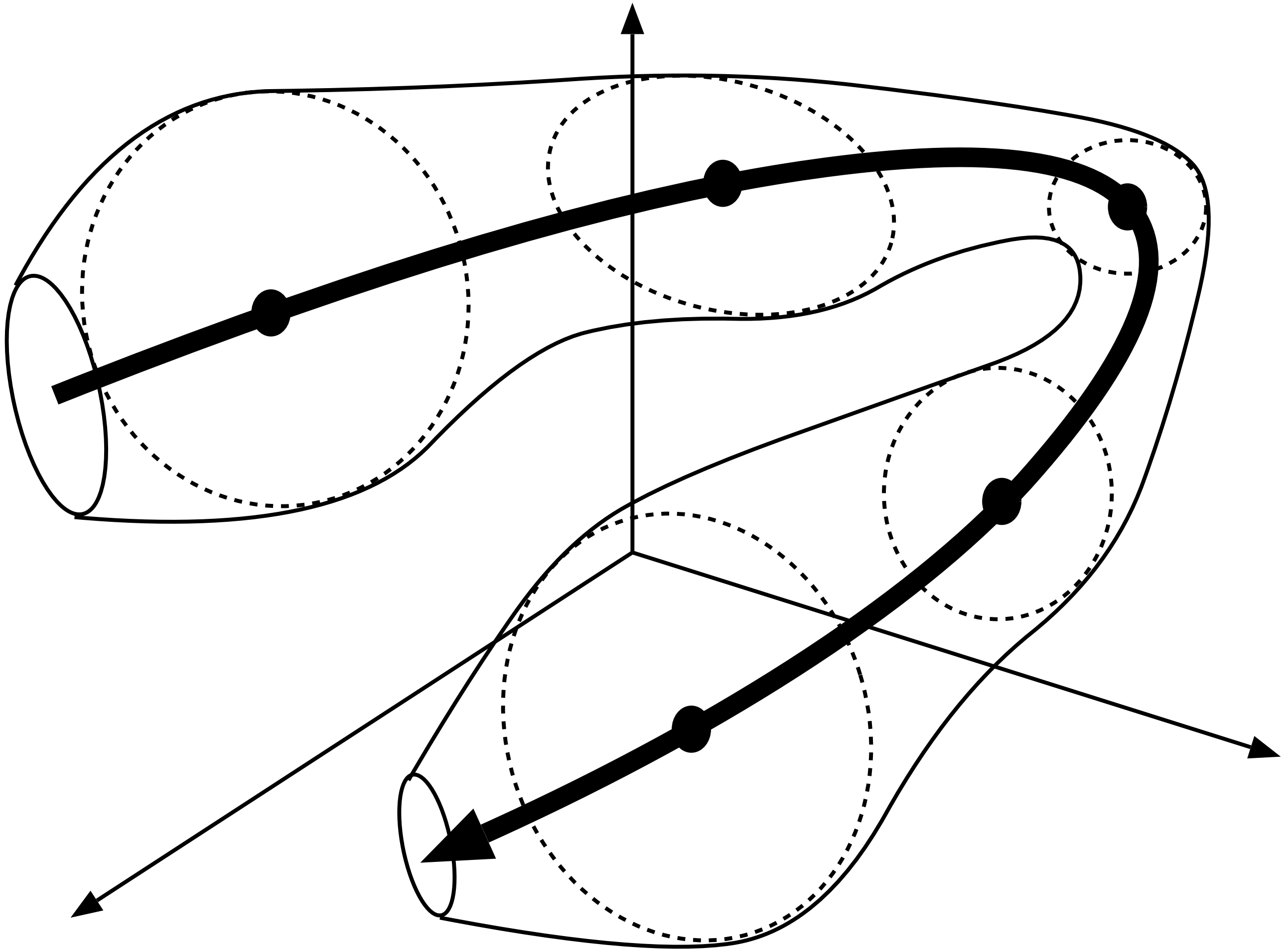

図6. 変動の大小に応じて管の太さを場所によって変える。すなわちマハラノビス距離などを使用する。そうすれば時間的にも粗くできる。これがHMMであると理解して良い。

通常の k-means クラスタリング

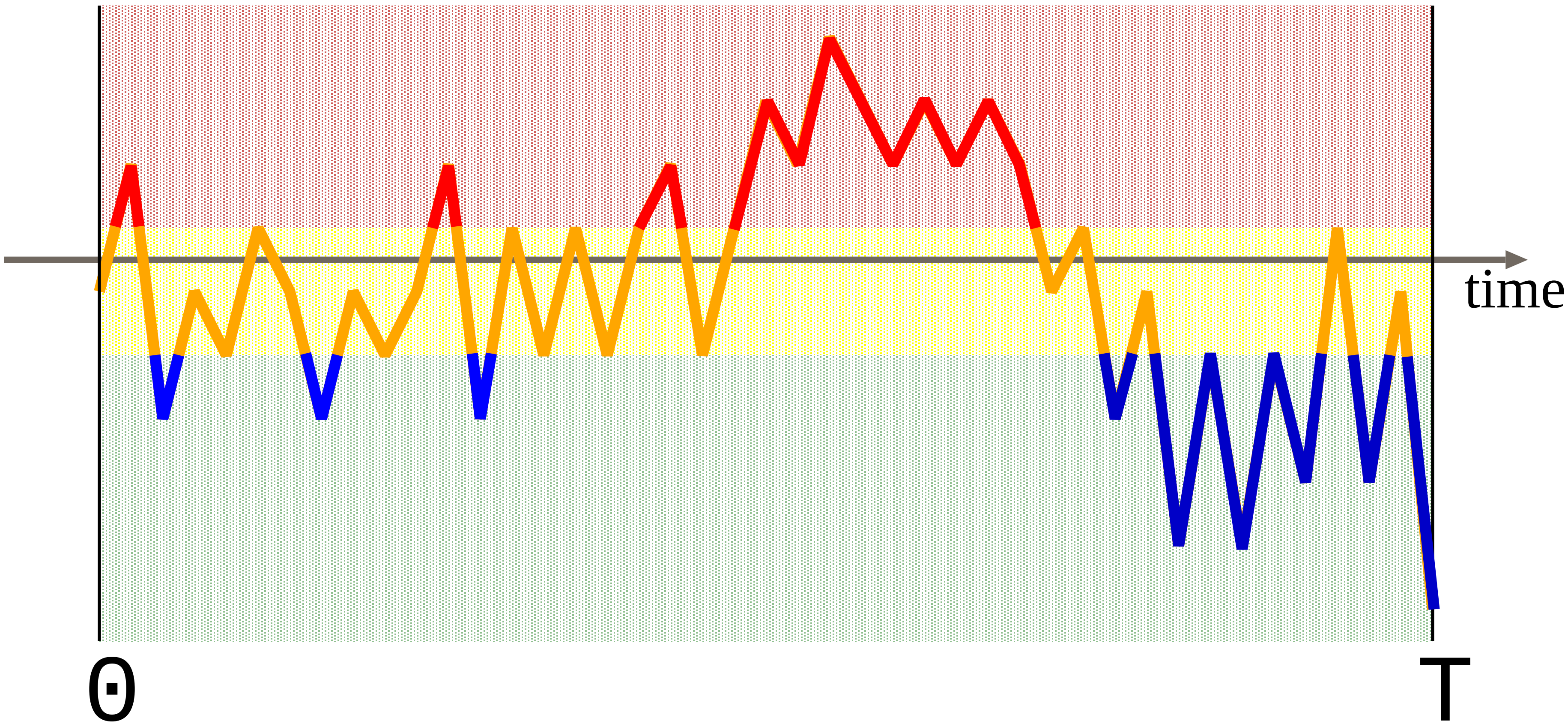

図 7. 全区間の値の集合を k-means アルゴリズムによりクラスタリングする。二乗距離の総和を最小とする。

Segmental k-means クラスタリング

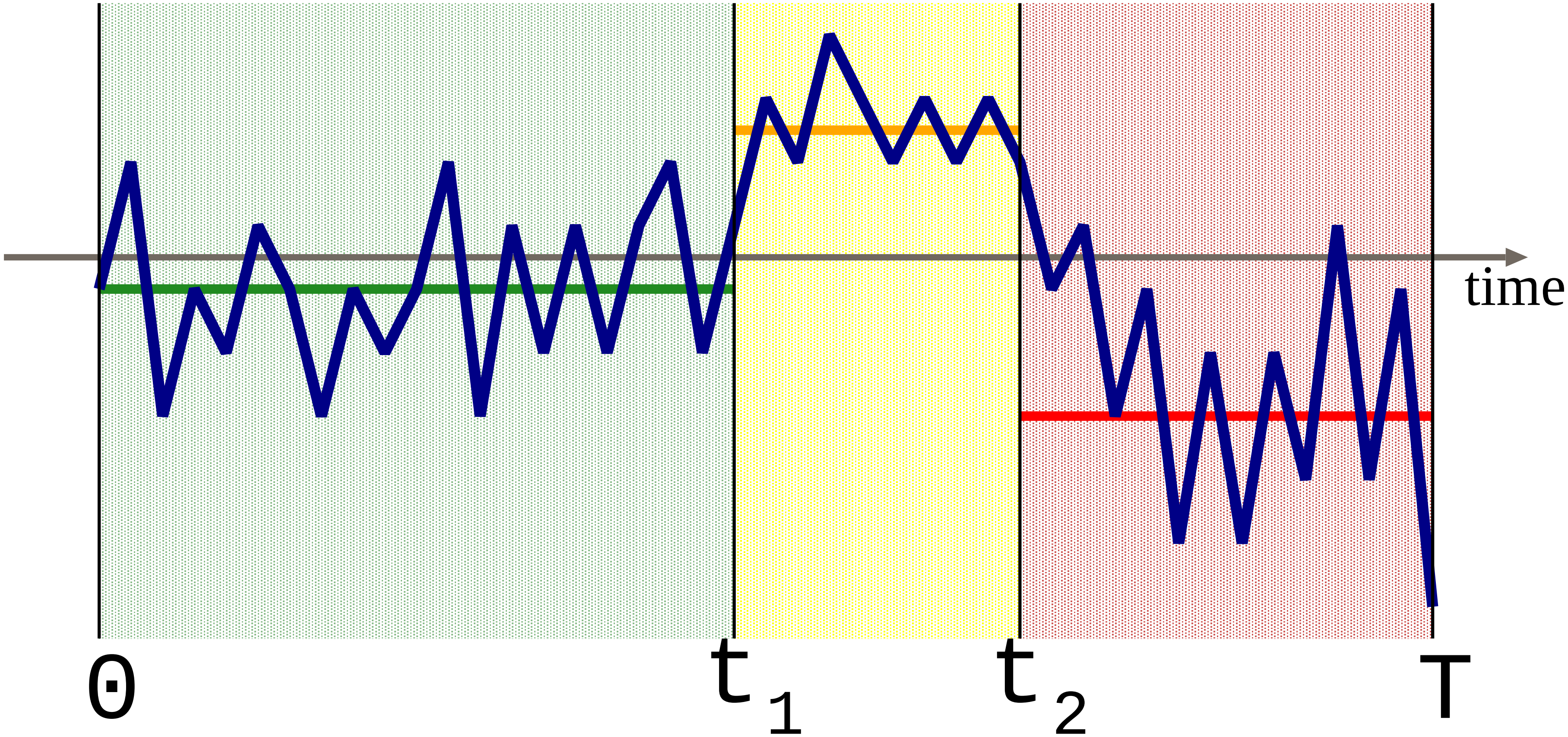

図 8. K **個の区間を持つ階段関数で置き換える。距離の二乗総和を最小とする。** t_1, t_2 **はどのように求めるか。**

Segmental k-means クラスタリング

- Euclid 距離二乗和最小化の segmental k-means クラスタリング：

$$D(t_1, t_2, \cdots, t_{K-1}) = \sum_{k=1}^{K} \sum_{t=t_{k-1}}^{t_k - 1} (x_t - m_k)^2$$

ただし $m_k = \frac{1}{t_k - t_{k-1}} \sum_{t=t_{k-1}}^{t_k - 1} x_t, \quad t_0 \equiv 0,\ t_K \equiv T$

をセグメント境界 $t_1, t_2, \cdots, t_{K-1}$ に関して最小にする。

- Segmental k-means アルゴリズム (二段階反復最小化アルゴリズム)
  1. 適当にセグメント境界初期値 $t_1, t_2, \cdots, t_{K-1}$ を与える。
  2. 各区間 $[t_{k-1}, t_k)$ 内で、平均値 $m_k = \frac{1}{t_k - t_{k-1}} \sum_{t=t_{k-1}}^{t_k - 1} x_t$ を算出する。(これは Euclid 距離二乗和を最小にするように m_k を決めること)
  3. 得られた平均値列 $\{m_1, m_2, \cdots, m_K\}$ と入力との間で DP (dynamic programming) マッチングして、Euclid 距離二乗和最小の経路を求めることにより、セグメント境界 $t_1, t_2, \cdots, t_{K-1}$ を更新する。
  4. 収束するまで、2 に戻って繰り返す。

平均値の意味

- サンプル値集合 $\{x_1, x_2, x_3, \cdots, x_N\}$ に対して、
  Euclid 距離二乗和 $D = \sum_{i=0}^{N-1}(x_i - m)^2$ を最小にする m は、
  $m = \frac{1}{N}\sum_{i=0}^{N-1} x_i$ (算術平均) である。
- 導出:
  $\frac{dD}{dm} = -2\sum_{i=0}^{N-1}(x_i - m) = 0$ と置くことにより、$\sum_{i=0}^{N-1} x_i = Nm$
  すなわち、$m = \frac{1}{N}\sum_{i=0}^{N-1} x_i$ を得る。

Segmental k-means クラスタリング 2

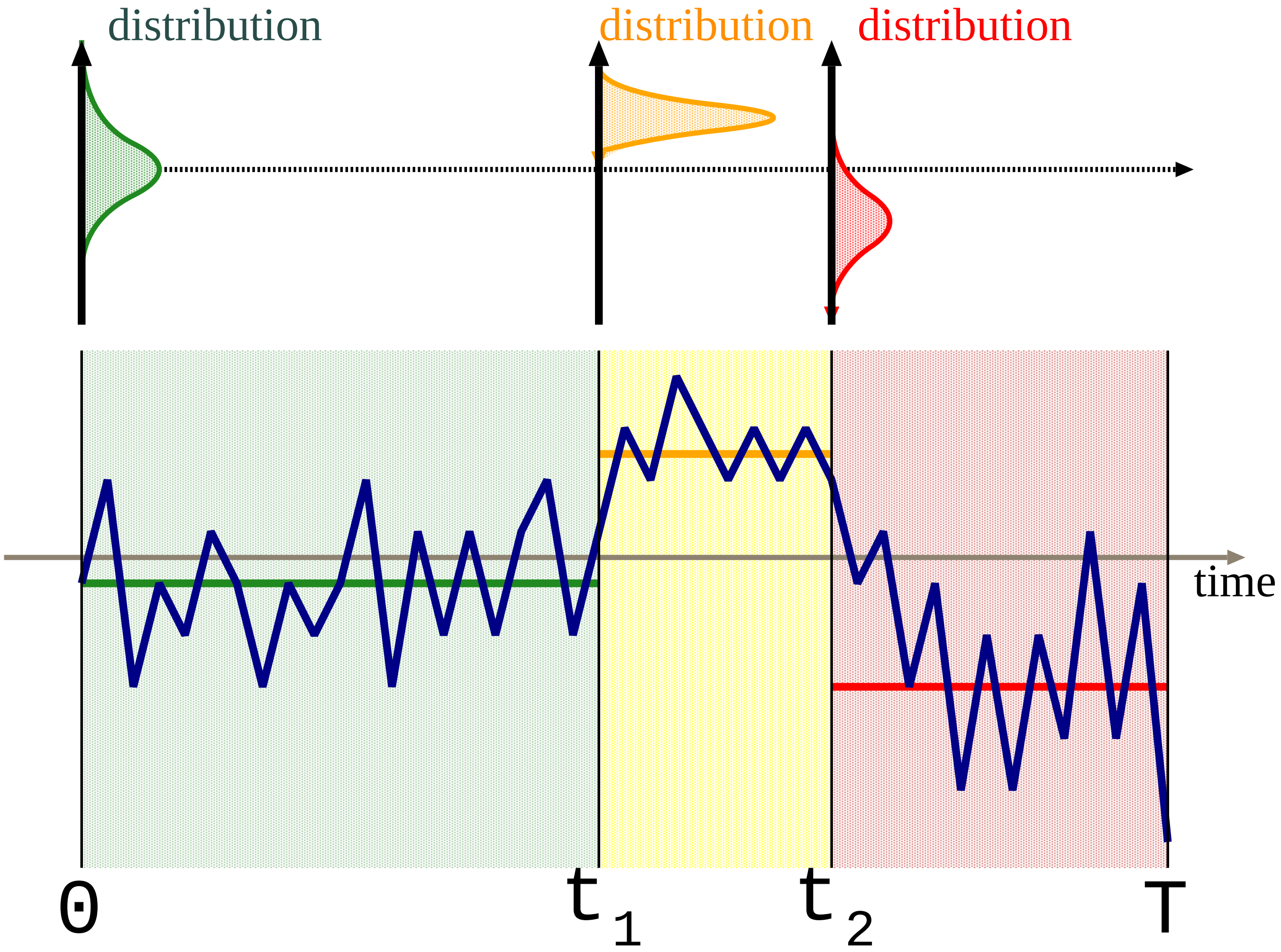

図9. 距離がマハラノビス距離で定義されている場合、距離の二乗総和を最小とするようにt_1, t_2を決定する。

Segmental k-means クラスタリング 2

- **Mahalanobis 距離和最小化の segmental k-means クラスタリング：**

$$D(t_1, t_2, \cdots, t_{K-1}) = \sum_{k=1}^{K} \sum_{t=t_{k-1}}^{t_k-1} \frac{(x_t - m_k)^2}{s_k^2}, \qquad t_0 \equiv 0,\ t_K \equiv T$$

ただし $m_k = \frac{1}{t_k - t_{k-1}} \sum_{t=t_{k-1}}^{t_k-1} x_t, \quad s_k^2 = \frac{1}{t_k - t_{k-1}} \sum_{t=t_{k-1}}^{t_k-1} (x_t - m_k)^2$

をセグメント境界 $t_1, t_2, \cdots, t_{K-1}$ **に関して最小にする。**

- **Segmental k-means アルゴリズム (二段階反復最小化アルゴリズム)**

1. **適当にセグメント境界初期値**$t_1, t_2, \cdots, t_{K-1}$**を与える。**
2. **各区間**$[t_{k-1}, t_k)$**内で、標本平均**$m_k = \frac{1}{t_k - t_{k-1}} \sum_{t=t_{k-1}}^{t_k-1} x_t$ **と標本分散**$s_k^2 = \frac{1}{t_k - t_{k-1}} \sum_{t=t_{k-1}}^{t_k-1}$ $m_k)^2$ **を算出する。(これはMahalanobis 距離和を最小にするように**m_k, s_k **を決めること)**
3. **得られた平均値列**$\{m_1, m_2, \cdots, m_K\}$**と入力との間で DP (dynamic programming) マッチングして、分散**s_k **に関する Mahalanobis 距離和最小の経路を求めることにより、セグメント境界**$t_1, t_2, \cdots, t_{K-1}$ **を更新する。**
4. **収束するまで、2に戻って繰り返す。**