Markov Models in Healthcare

Farzad Kamalzadeh
Health Care Problem

• Chronic diseases: A **chronic disease** is a human health condition or disease that is persistent or otherwise long-lasting in its effects or a disease that comes with time.

• *Chronic*: the course of the disease lasts for more than 3 months.

• Common chronic diseases include:
  • arthritis
  • Asthma
  • Cancer
  • heart failure
  • diabetes
  • hepatitis C
  • HIV/AIDS
Epidemiology

• Chronic diseases constitute a major cause of mortality:
  • WHO: 38 million deaths a year to non-communicable diseases
  • United States: 25% of adults have at least two chronic conditions
  • 1 in 2 Americans (133 million) has at least one chronic medical condition
  • 61% of all deaths among people older than 65 in the population

• Diabetes:
  • 7th leading cause of death in the US
  • Leading cause of many complications such as kidney failure, non-traumatic lower limb amputations, blindness
  • Major cause of heart disease
Economic impact

• Chronic diseases constitute a major section of medical care spending: (direct costs)
  • 75% of the $2 trillion spent annually in US medical care ($1.5 trillion)
  • Diabetes: $1 in $3 Medicare expenditure

• (indirect costs)
  • limitations in daily activities
  • loss in productivity
  • loss of days of work

• Diabetes: $322 billion per year
Nature of Chronic Diseases

Disease starts
-10
Complications start
-5
Symptoms
Disease severity
Medical intervention and treatments
Diagnosis
0
time

almost no symptoms
Natural Disease Progression

- **Asymptomatic stage**
- **Progressive disease (stage 1)**
- **Progressive disease (stage 2)**
- **Progressive disease (stage 3)**
- **Death**

Markov Chain
Markov Models

\[ (0) \text{Apparenty Healthy} \xrightarrow{\lambda_1} (1) \text{Asymptomatic Phase} \xrightarrow{\lambda_2} (2) \text{Symptomatic Phase} \xrightarrow{\lambda_3} \text{Death} \]

\[ (0) \text{Apparenty Healthy} \xrightarrow{\lambda_1} (1) \text{Asymptomatic Phase} \xrightarrow{\lambda_2} (2) \text{Symptomatic Phase} \xrightarrow{\lambda_3} (3) \text{Death from this disease} \]

\[ (4) \text{Other causes of death} \xrightarrow{\mu_1} (1) \text{Asymptomatic Phase} \xrightarrow{\mu_2} (2) \text{Symptomatic Phase} \xrightarrow{\mu_3} (3) \text{Death from this disease} \]
Markov Models

\[ P_{11} \quad P_{12} \quad P_{13} \quad P_{22} \quad P_{23} \quad P_{24} \quad P_{33} \quad P_{34} \quad P_{44} \]

States:
- S1: Vulnerable People
- S2: HIV Diagnosis
- S3: AIDS Diagnosis
- S4: Death

Transitions:
- \( P_{11} \): Stay in S1
- \( P_{12} \): From S1 to S2 (HIV Diagnosis)
- \( P_{13} \): From S1 to S3 (AIDS Diagnosis)
- \( P_{22} \): Stay in S2
- \( P_{23} \): From S2 to S3 (AIDS Diagnosis)
- \( P_{24} \): From S2 to S4 (Death)
- \( P_{33} \): Stay in S3
- \( P_{34} \): From S3 to S4 (Death)
- \( P_{44} \): Stay in S4
Markov Models
Markov Models
Markov Models
Markov Models
Markov Models

Population at breast cancer risk
taking part in either MSP (70%), OS (20%) or no screening
Markov Chains
Data Implementation

Asymptomatic stage

Progressive disease (stage 1) -> Progressive disease (stage 2) -> Progressive disease (stage 3) -> Death

\[ \lambda_1 = ? \]
\[ \lambda_2 = ? \]
\[ \lambda_3 = ? \]
\[ \lambda_4 = ? \]
\[ \lambda_5 = ? \]
\[ \lambda_6 = ? \]
\[ \lambda_7 = ? \]
Data Implementation

Table:

<table>
<thead>
<tr>
<th>Patient</th>
<th>t=1</th>
<th>t=2</th>
<th>t=3</th>
<th>t=4</th>
<th>t=5</th>
<th>t=6</th>
<th>t=7</th>
<th>t=8</th>
<th>t=9</th>
<th>t=10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 2</td>
<td>C</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 3</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td>A</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 4</td>
<td>B</td>
<td></td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Patient 5</td>
<td>C</td>
<td></td>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E</td>
</tr>
<tr>
<td>Patient 6</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
<td></td>
<td>E</td>
</tr>
<tr>
<td>Patient 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
<td>B</td>
</tr>
</tbody>
</table>
Hidden Markov Models

\[ N \]
\[ T \]
\[ \theta_{i=1\ldots N} \]
\[ \phi_{i=1\ldots N, j=1\ldots N} \]
\[ \phi_{i=1\ldots N} \]
\[ x_{t=1\ldots T} \]
\[ y_{t=1\ldots T} \]
\[ F(y|\theta) \]
\[ x_{t=2\ldots T} \]
\[ y_{t=1\ldots T} \sim \text{Categorical}(\phi_{x_{t-1}}) \]
\[ \sim F(\theta_{x_{t}}) \]
Hidden Markov Models - Learning

• The parameter learning task in HMMs: given an output sequence or a set of such sequences ===> the best set of state transition probabilities.

• The task is usually to derive the maximum likelihood estimate of the parameters of the HMM given the set of output sequences

• Local maximum likelihood can be derived efficiently using the Baum–Welch algorithm
Baum–Welch algorithm

\[ \lambda = (A, B, \pi) \]

for each sequence

while desired level of convergence not acquired

for \( t = 1 \) to \( T \)

for \( i \) in \( S \)

\[ \alpha_i(t) = P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_t = y_t | X_t = i, \lambda) \]

the probability of seeing the \( Y_1 = y_1, Y_2 = y_2, \ldots, Y_t = y_t \) and being in state \( i \) at time \( t \)

\[ \beta_i(t) = P(Y_{t+1} = y_{t+1}, Y_{t+2} = y_{t+2}, \ldots, Y_T = y_T | X_t = i, \lambda) \]

the probability of the ending partial sequence \( Y_{t+1} = y_{t+1}, Y_{t+2} = y_{t+2}, \ldots, Y_T = y_T \) given starting state \( i \) at time \( t \)

\[ \gamma_i(t) = P(X_t = i | Y, \lambda) = \frac{\alpha_i(t) \beta_i(t)}{\sum_{j=1}^{N} \alpha_j(t) \beta_j(t)} \]

the probability of being in state \( i \) at time \( t \) given the observed sequence \( Y \) and the parameters \( \lambda \)

\[ \delta_{ij}(t) = P(X_t = i, X_{t+1} = j | Y, \lambda) = \frac{\alpha_i(t) a_{ij} \beta_i(t + 1) b_j(y_{t+1})}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i(t) a_{ij} \beta_i(t + 1) b_j(y_{t+1})} \]

the probability of being in state \( i \) and \( j \) at times \( t \) and \( t+1 \) respectively given the observed sequence \( Y \) and parameters \( \lambda \)

\[
\text{update: } \quad \pi_i = \gamma_i(1) \quad a_{ij} = \frac{\sum_{t=1}^{T-1} \delta_{ij}(t)}{\sum_{t=1}^{T} \gamma_i(t)} \quad b_i(v_k) = \frac{\sum_{t=1}^{T} 1_{y_t=v_k} \gamma_i(t)}{\sum_{t=1}^{T} \gamma_i(t)}
\]
Baum–Welch algorithm
References


