Ugeopgave 4, Mønstergenkendelse

Søren Hauberg (hauberg@diku.dk)

10. marts 2007

Exercise 8.14

Vi ser på fejlfunktionen

$$E(\mathbf{x}, \mathbf{y}) = h \sum_i x_i - \beta \sum_{i,j} x_i x_j - \eta \sum_i x_i y_i \tag{1}$$

hvor $x_i \in \{-1, 1\}$ og $y_i \in \{-1, +1\}$. I denne opgave er $\beta = h = 0$ hvilket betyder at E reduceres til

$$E(\mathbf{x}, \mathbf{y}) = -\eta \sum_i x_i y_i \qquad \eta > 0 \tag{2}$$

Vi ved at

$$p(\mathbf{x}, \mathbf{y}) = \frac{1}{Z} \exp\left[-E(\mathbf{x}, \mathbf{y})\right] = \frac{1}{Z} \exp\left[\eta \sum_i x_i y_i\right]. \tag{3}$$

For at maksimere p skal vi altså maksimere $\sum_i x_i y_i$. Denne sum antager maksimal værdi når $x_i y_i$ er maksimal for alle i. Det ses trivielt at

$$\begin{aligned}
x_i = +1 \wedge y_i = +1 \quad &\Rightarrow \quad x_i y_i = +1 &(4)\\
x_i = -1 \wedge y_i = +1 \quad &\Rightarrow \quad x_i y_i = -1 &(5)\\
x_i = +1 \wedge y_i = -1 \quad &\Rightarrow \quad x_i y_i = -1 &(6)\\
x_i = -1 \wedge y_i = -1 \quad &\Rightarrow \quad x_i y_i = +1 &(7)
\end{aligned}$$

Altså er $x_i y_i$ maksimal når $(x_i = -1 \wedge y_i = -1) \vee (x_i = +1 \wedge y_i = +1)$. Dette er det sammen som $x_i = y_i$ hvorved det ønskede er vist.

Computer Exercise

Vi ser på en diskret stokastisk variabel $X_n \in \{regn, solskin\}$. Vi antager at

$$p(X_{n+1} | \{X_i\}_{i=1..n}) = p(X_{n+1} | X_n) \tag{8}$$

altså en almindelig første-ordens Markov antagelse. Det antages desuden at

$$\begin{aligned}
p(X_{n+1} = regn | X_n = regn) &= 0.8 &(9)\\
p(X_{n+1} = solskin | X_n = regn) &= 0.2 &(10)\\
p(X_{n+1} = regn | X_n = solskin) &= 0.1 &(11)\\
p(X_{n+1} = solskin | X_n = solskin) &= 0.9 &(12)
\end{aligned}$$

Disse sandsynligheder opstilles i overgangsmatricen T

$$T = \left(\begin{array}{cc} 0.8 & 0.2 \\ 0.1 & 0.9 \end{array} \right). \tag{13}$$

Til slut antages det at

$$p(X_1 = regn) = 0.2 \qquad \text{og} \qquad p(X_1 = solskin) = 0.8 \tag{14}$$

Det ses at

$$\begin{aligned} p(X_2) &= \sum_{X_1 \in \{regn, solskin\}} p(X_2, X_1) && (15) \\ &= \sum_{X_1 \in \{regn, solskin\}} p(X_2 | X_1) p(X_1) && (16) \\ &= T^T p(X_1). && (17) \end{aligned}$$

Dette genereliserer trivielt til

$$p(X_{n+1}) = T^T p(X_n) \tag{18}$$

På figur 1 ses sandsynlighederne for *regn* og *solskin* som funktion af n.

Som det ses konvergerer sandsynlighederne, hvilket vil sige at

$$p(X_{n+1}) = p(X_n). \tag{19}$$

Indsættes resultatet fra tidligere fås

$$T^T p(X_n) = p(X_n) \tag{20}$$

hvilket er et egenværdi problem hvor egenværdien er 1. Dvs. at $p(X_n)$ ved konvergens er den egenvektor af T^T som har egenværdien 1. Disse værdier er vist med grønt på figuren.

Kildekode til GNU Octave

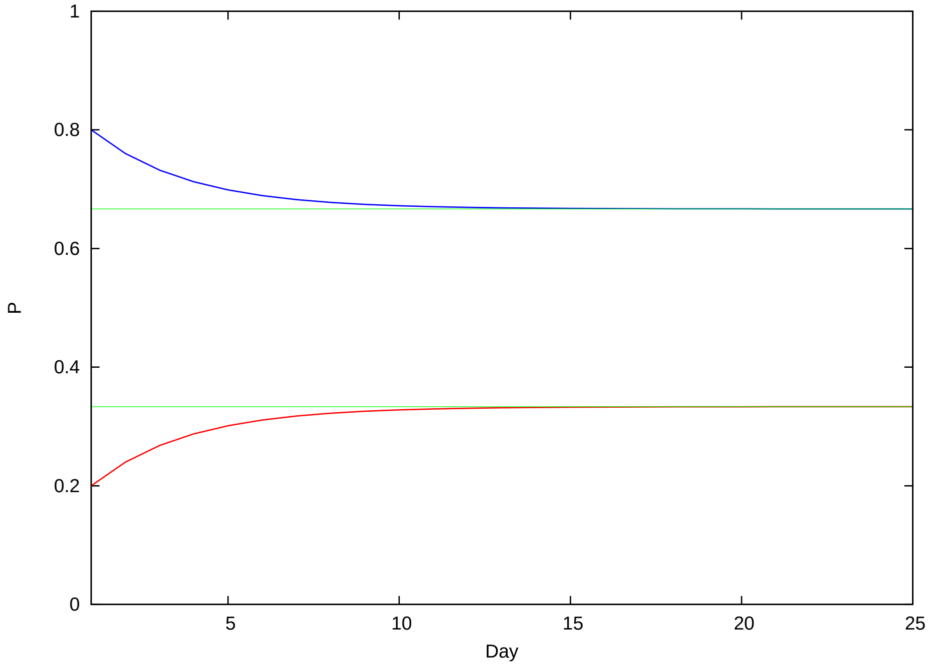
```
clear all

## MODEL
## We're working with a discrete stochastic variable X_n \in {rain, sun}.
## It is assumed that
##   p(X_{n+1} | {X_i}_{i = 1..n}) = p(X_{n+1} | X_n)
## i.e. a standard first-order Markov assumption.
##

## GIVEN PROBABILITIES
##              |   rain_{n+1}   |    sun_{n+1}
##    ----------+----------------+----------------
##    rain_n    |      0.8       |       0.2
##    sun_n     |      0.1       |       0.9
##
T = [ 0.8, 0.2;
      0.1, 0.9 ];

## PRIOR
##   p(rain_1) = 0.2  and  p(sun_1) = 0.8
##
prior = [ 0.2; 0.8 ];

## COMPUTE THE PROBABILITY OF RAIN ON DAY n
N = 25;
P = zeros(2,N); P(:,1) = prior;
for i = 2:N
  P(:,i) = T' * P(:,i-1);
endfor
```



Figur 1: Resultat af at afvikle programmet. Den blå kurve angiver sandsynligheden for solskin, mens den røde angiver for regn. De grønne kurver angiver de asymptotiske sandsynligheder.

```
## WHAT WILL THE PROBABILITIES CONVERGE TO?
##  Convergense is when
##     p(X_{n+1}) = p(X_n)
##  which is the same as
##     T' * p(X_n) = p(X_n)
##  If we solve this eigen equation and find the eigen vector
##  with an eigen value of 1 (in the equation the lambda on the right is 1)
##  we have p(X_n).
[V, E] = eig(T');
i = find(diag(E) == 1);
Pc = V(:,i)/sum(V(:,i));

## Plot
figure(1), clearplot
  h1 = line(1:N, P(1,:));
  h2 = line(1:N, P(2,:));
  h3 = line([1,N], [Pc(1), Pc(1)]);
  h4 = line([1,N], [Pc(2), Pc(2)]);
  set(h1, "color", [1,0,0], "linewidth", 2);
  set(h2, "color", [0,0,1], "linewidth", 2);
  set(h3, "color", [0,1,0], "linewidth", 1);
  set(h4, "color", [0,1,0], "linewidth", 1);
  xlabel("Day")
  ylabel("P")
  axis([1, N, 0, 1])
```