Midterm Exam

Prof. Jim Rehg
CS 7635 Computational Perception
College of Computing
Georgia Institute of Technology

March 13, 2002

The exam is take-home. You may refer to any notes, books, etc. but everyone must do their own work. Please hand in your solutions by 5pm on Thursday, March 14.

1. Face Recognition [25 points]
What makes face recognition difficult? List three factors in order of decreasing importance and discuss each one briefly. Justify your choice of order. What do you think is the most promising direction for face recognition research and why?

2. PCA [25 points]
A standard technique, which is sometimes called the "SVD trick," makes it possible to dramatically reduce the computational cost of PCA when the number of example vectors is much smaller than the dimension of the vector space.

Let $\mathbf{Y}$ be the standard n by p data matrix, containing p data vectors of dimension n, where $p \ll n$. The *left singular vectors* $\{\mathbf{u}_i\}$ of $\mathbf{Y}$ are orthonormal vectors satisfying $\mathbf{Y}\mathbf{Y}^T\mathbf{u}_i = \sigma_i^2\mathbf{u}_i$. Similarly, the orthonormal set of right singular vectors $\{\mathbf{v}_i\}$ satisfy $\mathbf{Y}^T\mathbf{Y}\mathbf{v}_i = \lambda_i^2\mathbf{v}_i$.

(a) Compute $\|\mathbf{Y}^T\mathbf{u}_i\|$ and $\|\mathbf{Y}\mathbf{v}_i\|$.

(b) Show that the left and right singular vectors satisfy:

$$\mathbf{Y}\mathbf{v}_i = \sigma_i\mathbf{u}_i$$
$$\mathbf{Y}^T\mathbf{u}_i = \lambda_i\mathbf{v}_i$$
$$\sigma_i = \lambda_i$$

(c) Let $\mathbf{S}$ denote the usual sample covariance matrix. Using the result from (b), give a complete recipe for computing the principle components of $\mathbf{S}$ in much less time than it would take to perform an eigenanalysis of $\mathbf{S}$.

(d) Extra Credit +5: How are the equations above related to the Singular Value Decomposition (SVD)?

3. Markov Chains [15 points]
A Markov model is called *irreducible* if it is possible to get from any one hidden state to any other hidden state in a finite number of steps with nonzero probability.[1] Determine whether each of the following Markov models is irreducible and explain your answer.

$$(1) \qquad \Pi = \begin{bmatrix} 0.5 & 0.5 & 0.0 & 0.0 \\ 0.9 & 0.1 & 0.0 & 0.0 \\ 0.0 & 0.2 & 0.5 & 0.3 \\ 0.1 & 0.0 & 0.0 & 0.9 \end{bmatrix}$$

$$(2) \qquad \Pi = \begin{bmatrix} 0.0 & 0.8 & 0.0 & 0.2 \\ 0.0 & 0.0 & 0.2 & 0.8 \\ 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 \end{bmatrix}$$

$$(3) \qquad \Pi = \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \\ 0.0 & 1.0 & 0.0 & 0.0 \end{bmatrix}$$

[1] By a "step" we mean one tick of the clock in which a state transition is chosen at random according to the transition probabilities and the state machine is advanced to the next state.

4. Markov Models [15 points]
(a) A high school student has three possible outcomes for each school year: dropping out with probability d, repeating that year with probability r, and advancing to the next year with probability a. There are four years of high school, after which the student graduates. Develop a Markov model for this probabilistic matriculation scenario. Define the state space and write the state transition matrix along with any constraints on the model parameters.

(b) Extra Credit +5 Let $p(s_{t-1})$ be the probability distribution over the state of an arbitrary Markov model at time $t - 1$. Write an expression for $p(s_t)$. In part (a), let $s = i$ denote the state of being in the ith year of high school. Compute $p(s_3 = 3)$ under the assumption that $p(s_1 = 1) = 1$ (start out as first year student).

5. Hidden Markov Models [20 points]
Parameter learning in HMM's can be expressed as the recursive application of the *forward-backward* equations:

$$\alpha_t(k) \equiv p(\mathbf{Y}_t, s_t = k) \quad \text{(Forward Recursion)}$$
$$\alpha_t(k) = p(y_t|s_t = k) \sum_{j=1}^{n} p(s_t = k|s_{t-1} = j)\alpha_{t-1}(j)$$
$$\alpha_0(k) = p(s_0 = k)$$

$$\beta_t(k) \equiv p(\bar{\mathbf{Y}}_{t+1}|s_t = k) \quad \text{(Backward Recursion)}$$
$$\beta_t(k) = \sum_{j=1}^{n} p(s_{t+1} = j|s_t = k)p(y_{t+1}|s_{t+1} = j)\beta_{t+1}(j)$$
$$\beta_T(k) = 1,$$

where $\mathbf{Y}_t = \{y_1, y_2, \ldots, y_t\}$ and $\bar{\mathbf{Y}}_{t+1} = \{y_{t+1}, y_{t+2}, \ldots, y_T\}$.

(a) Show that both the forward and backward recursions can be used (separately) to compute $p(\mathbf{Y}_T)$.

(b) Define $\gamma_t(k) = p(s_t = k|\mathbf{Y}_T)$. Show that

$$\gamma_t(k) = \frac{\alpha_t(k)\beta_t(k)}{\sum_{j=1}^{n} \alpha_t(j)\beta_t(j)}$$