HW 5: Learning Theory II (580.692)

Instructor: René Vidal, Office: 308B Clark, E-mail: *rvidal@cis.jhu.edu*
Grader: Dheeraj Singaraju. Office: 319 Clark, E-mail: *dheeraj@cis.jhu.edu*

Due 10/26/06 beginning of the class

1. Read Chapter 4 of GPCA book. Go to *http://www.vision.jhu.edu/gpcabook/* and submit all the typos you find as well as suggestions you may have to improve the quality and/or readability of the material. **You will receive credit for each interesting typo or suggestion you submit**.

2. **Hyperplane clustering with correspondences**: Exercise 4.5 of the book.

3. **Clustering linear from bilinear varieties**: Let $S = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^D \times \mathbb{R}^D : \boldsymbol{u}^\top \boldsymbol{x} = 0 \vee \boldsymbol{y}^\top A \boldsymbol{x} = 0\}$, where $\boldsymbol{u} \neq \boldsymbol{0}$ and $e_D^\top A = e_D^\top = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$.

   (a) Find a polynomial $p(\boldsymbol{x}, \boldsymbol{y})$ that vanishes on S. How many independent monomials are there in p?

   (b) Show that p can be written as $p(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{y}^\top \mathcal{M} \nu_2(\boldsymbol{x})$. How is $\mathcal{M}$ related to A and $\boldsymbol{u}$?

   (c) If $D = 3$, write the $\mathcal{M}$ explicitly, and show how one can compute $\boldsymbol{u}$ and A from the entries of $\mathcal{M}$.

   (d) Let $\boldsymbol{X} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i) \in S\}_{i=1}^N$ be a given data set. Derive an algorithm to compute $\mathcal{M}$ from $\boldsymbol{X}$. Then, derive an algorithm to compute A and $\boldsymbol{u}$ from the derivatives of p with respect to $\boldsymbol{x}$ and $\boldsymbol{y}$.
   Hint: *Explicitly write down the derivatives of the polynomial $p(\boldsymbol{x}, \boldsymbol{y})$ and inspect the values of these derivatives when $\boldsymbol{u}^\top \boldsymbol{x} = 0$ and when $\boldsymbol{y}^\top A \boldsymbol{x} = 0$. Also, making clever canonical choices for $\boldsymbol{x}$ and $\boldsymbol{y}$ can make the solution easier.*

4. **Clustering bilinear varieties**. Let $S_j = \{(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \boldsymbol{y}^\top F_j \boldsymbol{x} = 0\}$ for $j = 1, 2$. Assume that for $j = 1, 2$ there are nonzero vectors $\boldsymbol{e}_1 \neq \boldsymbol{e}_2$ such that $\boldsymbol{e}_j^\top F_j = \boldsymbol{0}^\top$.

   (a) Show that $\boldsymbol{y}^\top A \boldsymbol{x} = (\boldsymbol{y} \otimes \boldsymbol{x})^\top \mathrm{vec}(A)$, where $\mathrm{vec}(A)$ is the stack of all the rows of A.

   (b) Let $\boldsymbol{z} = \boldsymbol{y} \otimes \boldsymbol{x}$. Find a polynomial $p(\boldsymbol{z}) = \boldsymbol{c}^\top \nu_2(\boldsymbol{z})$ that vanishes on $S_1 \cup S_2$. How many independent monomials in $\boldsymbol{z}$ are there in $p(\boldsymbol{z})$?

   (c) Find a polynomial $p(\boldsymbol{x}, \boldsymbol{y}) = \nu_2(\boldsymbol{y})^\top \mathcal{F} \nu_2(\boldsymbol{x})$ that vanishes on $S_1 \cup S_2$. How is $\mathcal{F}$ related to F_1 and F_2? How many independent monomials in $(\boldsymbol{x}, \boldsymbol{y})$ are there in $p(\boldsymbol{x}, \boldsymbol{y})$? Why does this number not coincide with the one computed in part (b)?

   (d) Given $\boldsymbol{X} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i) \in S_1 \cup S_2\}_{i=1}^N$, derive an algorithm for computing $\mathcal{F}$ from $\boldsymbol{X}$ using parts (a) & (c).

   (e) Show that for all $(\boldsymbol{x}, \boldsymbol{y}) \in S_j$, we have $\boldsymbol{e}_j^\top \frac{\partial p}{\partial \boldsymbol{y}}(\boldsymbol{x}, \boldsymbol{y}) = 0$.

   (f) Given $\boldsymbol{X}$, let $\boldsymbol{\ell}_i = \frac{\partial p}{\partial \boldsymbol{y}}(\boldsymbol{x}_i, \boldsymbol{y}_i)$. Use part (e) to show that $\{\boldsymbol{\ell}_i\}_{i=1}^N$ live in a union of two planes with normals $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$. How would you compute $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ from $\{\boldsymbol{\ell}_i\}_{i=1}^N$ and cluster the data $\boldsymbol{X}$ into the two groups?

5. Implement the `GPCA` algorithm for hyperplanes. Test it on the datasets on the website. Use the following format

Function `[b,group] = gpca(x,n)`

Parameters

`x`	$D \times N$ matrix whose columns are the data points
`n`	number of groups

Returned values

`b`	$D \times n$ matrix whose columns are the normal vectors
`group`	$1 \times N$ vector with group membership of each point

Description

Hyperplane clustering using GPCA.