Assignment 7

Due June 7 at 11:59PM . No late submissions allowed.

Please answer all questions. Students are expected to solve the problems on their own. No discussions allowed. The majority of the grading will be based on whether you are approaching the problem in the right way, even if you are not able to completely answer the question. Copying in any form will be dealt with very seriously, including possible reporting to campus committees.

1. (2 pts.) Given training data $\{x_i\}_{i=1}^L$, we want to learn the parameters $\theta = \{\lambda_k, \mu_k, \Sigma_k\}_{k=1}^K$ of a Mixture of Gaussians model using the EM algorithm. Derive the update rules for θ.

2. (2 pts.) Consider the multiclass logistic regressions problem in Sec. 4.3.4 of Pattern Recognition and Machine Learning by Bishop. It provides an outline of an iterative algorithm for the multi-class classification problem using the softmax function. Complete the derivation of (4.110) by filling in the details in the outline.

3. (2 pts.) Let $(\tilde{x}, \tilde{y})$ represent the image plane coordinates, whose corresponding representation in homogeneous coordinates is (x, y, z). Consider a camera model where $\tilde{x} = x$ and $\tilde{y} = \frac{y}{z}$. Prove that the projection of a 3D line on the image plane of this camera is a hyperbola.

4. (3 pts.)

   a) In the notation we have followed for stereo reconstruction, prove that $\mathbf{F} = \mathbf{e}_r \times \mathbf{H}$, where $\mathbf{p}_r = \mathbf{H}\mathbf{p}_l$.

   b) Consider the stereo system as defined in class, but with the world origin at the center of projection of the left camera and all distances represented with respect to this coordinate system. This implies $\mathbf{M}_l = \mathbf{M}_{l,int}\,[I_{3\times3}|\mathbf{0}]$ and $\mathbf{M}_r = \mathbf{M}_{r,int}\,[\mathbf{R}|\mathbf{t}]$. Prove that the fundamental matrix $\mathbf{F} = \hat{\mathbf{e}}_r\mathbf{M}_r\mathbf{M}_l^+$. What is the center of projection of the right camera in this coordinate system?

5. (3 pts.) Consider a point which moves along the straight line AC with constant velocity $V = [V_X, 0, V_Z]^T$, i.e. the motion is constrained in the XZ plane. The starting point A is denoted as $[X_{ref}, Y_{ref}, Z_{ref}]$, and AC is at an angle θ to AB which is parallel to the image plane. The image plane is the XY plane of the coordinate system, the principal point is the center of the image and the focal length of the camera is f. Assume perspective projection.

   a) Let $[x, y]^T$ be the projection of a 3D point $[X, Y, Z]^T$, along AC, on the image plane. Write down the differential equations relating the velocity of the $[x, y]^T$ to the 3D velocity and depth, for the particular motion for this point.

   b) If $\cot(\alpha) = \frac{\dot{x}}{\dot{y}}$, prove that

$$\cot(\theta) = \frac{1}{f}(x_{ref} - y_{ref}\cot(\alpha)), \tag{1}$$

   where the lower case letters represent the perspective projections of the upper case letters. (Hint: Use the constant velocity equation of motion to express the position of $[X, Y, Z]^T$ in terms of $[X_{ref}, Y_{ref}, Z_{ref}]$ and V. Then follow the perspective projection equations.)

   c) Write down the rotation matrix, in terms of θ, to rotate a point on AC to its corresponding location on AB.

   d) If $[X_\theta, Y_\theta, Z_\theta]^T$ represents a point on AC, and $[X_0, Y_0, Z_0]^T$ represents the corresponding point on AB, write down the geometrical relationship between these corresponding pair of 3D points.

Following some algebraic manipulation and using the equations of perspective projection and the assumption that $Z_\theta = Z_{ref}$ (this is actually an approximation), prove that

$$\begin{aligned} x_0 &= f\frac{x_\theta \cos(\theta) + x_{ref}(1-\cos(\theta))}{-\sin(\theta)(x_\theta + x_{ref}) + f} \\ y_0 &= \frac{y_\theta}{-\sin(\theta)(x_\theta + x_{ref}) + f}. \end{aligned}$$

6. (3 pts.) Consider the stereo system with the world origin at the center of projection of the left camera and all distances represented with respect to this coordinate system. This implies $\mathbf{M}_l = \mathbf{K}_l\,[I_{3\times 3}|\mathbf{0}]$ and $\mathbf{M}_r = \mathbf{K}_r\,[\mathbf{R}|\mathbf{t}]$, where $\mathbf{K}_l$ and $\mathbf{K}_r$ are the intrinsic parameters of the left and right camera, respectively. $(\mathbf{R}, \mathbf{t}) \in SE(3)$ represent the orientation and position of the right camera w.r.t. the left one. Recall that the fundamental matrix $\mathbf{F} = \hat{\mathbf{e}}_r\mathbf{M}_r\mathbf{M}_l^+$, where $\mathbf{M}_l^+$ is the pseudo-inverse of $\mathbf{M}_l$.
a) Prove that the left epipole $\mathbf{e}_l = \mathbf{K}_l\mathbf{R}^T\mathbf{t}$.
b) Prove that the right epipole $\mathbf{e}_r = \mathbf{K}_r\mathbf{t}$.
c) Prove that $\mathbf{M}_r\mathbf{M}_l^+ = \mathbf{K}_r\mathbf{R}\mathbf{K}_l^{-1}$.
d) Using the result $\hat{\mathbf{x}}\mathbf{S} = S^{-T}\left[\mathbf{S}^{-1}\mathbf{x}\right]_\times$ ($[\mathbf{w}]_\times$ is the same as $\hat{\mathbf{w}} \in SO(3)$), and the above expression of the fundamental matrix, show that $\mathbf{F} = \mathbf{K}_r^{-T}\mathbf{R}\mathbf{K}_l^T\left[\mathbf{K}_l\mathbf{R}^T\mathbf{t}\right]_\times$. (Hint: Express the fundamental matrix using the result in part (c). Apply the above result, three times, one matrix at a time, on the fundamental matrix.)
e) If the right camera is only translated w.r.t. the left camera (no rotation) and the intrinsic parameters of the two cameras are the same, prove that $\mathbf{F} = \hat{\mathbf{e}}_r$. If the translation is parallel to the x-axis and the matrix of intrinsic parameters is identity ($\mathbf{I}$), compute the fundamental matrix?