Statistical and Learning Techniques in Computer Vision
Homework 1: Due Thursday September 7, 2006

Practice Problems

The following problems are purely for practice. You should make sure you can do them (except perhaps for the last one) before continuing to the second set of problems, whose solutions are to be turned in for grading.

1. Derive the mean and the variance of the triangle distribution:

$$f(x) = \begin{cases} 0 & x < a \\ \frac{1}{2}(x-a) & a \le x \le a+1 \\ \frac{1}{2}(a+2-x) & a+1 < x \le a+2 \\ 0 & x > a+2 \end{cases}$$

2. Prove that the covariance of two independent random variables is 0. Explain intuitively why the converse is not necessarily true.

3. Given random variable $\mathbf{x}$ and random variable $\mathbf{y} = a\mathbf{x} + b$, derive the mean and variance of $\mathbf{y}$ as a function of the mean and variance of $\mathbf{x}$.

4. Repeat the previous exercise for the case that $\mathbf{x}$ and $\mathbf{y}$ are k-dimensional vectors, replacing a by $k \times k$ invertible matrix $\mathbf{A}$. Ignore b (i.e. assume it is 0).

5. Derive the maximum likelihood estimate of the parameter λ of the exponential:

$$f(x) = \lambda e^{\lambda x}.$$

6. Derive the maximum likelihood estimates of the mean and variance of the multivariate Gaussian. This will require a few advanced techniques in matrix manipulation, including derivatives with respect to determinants.

Homework Problems

Solutions to the following problems should be submitted for grading.

1. **(10 points)** Given a random vector $\mathbf{x}$ whose distribution is multivariate normal, i.e. $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, prove that when the covariance matrix, $\boldsymbol{\Sigma}$, is diagonal then the individual random variables that form the vector are independent.

2. **(10 points)** Consider the following smoothing (convolution) operator on 1d images, with pixel locations x_i:

$$x'_i = \frac{1}{2}x_i + \frac{1}{4}(x_{i-1} + x_{i+1}).$$

Under the assumption that all pixels in an image are independent and have the same mean, μ, and variance, σ^2, what are the mean and variance of x'_i?

3. **(20 points)** How good is the independence assumption stated in the previous problem? One way to address this question is to study the statistics of the difference between two pixels as a function of how far apart they are in the image. Here is one way to to this. For given distance δ, let

$$\mathcal{S}(\delta) = \{(\mathbf{x}, \mathbf{y}) \ \mid \ \|\mathbf{x} - \mathbf{y}\| = \delta\}$$

be the set of all pixel locations that are distance δ apart in the image, with the distance determined by the "city block" distance metric, i.e.

$$\|\mathbf{x} - \mathbf{y}\| = |x_2 - x_1| + |y_2 - y_1| \ .$$

Let $N(\delta)$ represent the number of pairs in $\mathcal{S}(\delta)$. Now we can compute

$$\mathrm{std}(\delta) = \left[\frac{\sum_{(\mathbf{x},\mathbf{y})\in\mathcal{S}} (I(\mathbf{x}) - I(\mathbf{y}))^2}{N(\delta)}\right]^{1/2} .$$

This is the standard deviation of the difference in intensity values as a function of the distance between pixels.

Write a Matlab program that takes an image and computes

$$\mathrm{std}(1), \mathrm{std}(2), \ldots, \mathrm{std}(25).$$

It should both output and plot these values. Test the example images posted on the course web page. Explain the results you obtain.

4. (**10 points**) Derive the MAP estimate for fitting a regression plane to a set of points $\mathcal{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$, with corresponding measurements $\mathbf{y}^\top = (y_1, \ldots, y_N)$. Assume the plane has the form

$$y = a_0 + \mathbf{a}^\top \mathbf{x}$$

and assume the prior distribution $p(\mathbf{a})$ is a normal distribution with mean $\mathbf{0}$ and covariance $\mathbf{\Sigma}$. Note that there is no prior on the intercept parameter a_0, which means it is not biased toward any particular location. You should assume that the measurement errors are normal with unknown variance σ^2 and estimate this variance as part of your solution.

5. (**15 points**) The expectation function can be treated as an operator on functions and on random variables. We can use this in a variety of analysis problems. The following explores this idea, using it to analyze the maximum likelihood estimator.

   (a) Prove that if $\mathbf{x}$ and $\mathbf{y}$ are random variables, with joint density $p(x, y)$, f and g and functions, and a and b are non-random values, then

   $$\mathbf{E}[af(\mathbf{x}) + bg(\mathbf{y})] = a\mathbf{E}[f(\mathbf{x})] + b\mathbf{E}[g(\mathbf{y})].$$

   One immediate result of this (which you do not need to prove) is that

   $$\mathbf{E}[\sum_{i=1}^{N} \mathbf{x}_i] = \sum_{i=1}^{N} \mathbf{E}[\mathbf{x}_i].$$

(b) Show that if $\mathbf{x}$ and $\mathbf{y}$ are independent then

$$\mathbf{E}[\mathbf{x}\mathbf{y}] = \mathbf{E}[\mathbf{x}]\mathbf{E}[\mathbf{y}].$$

(c) In the rest of the question we will use the above to prove properties of the MLE estimate $(\hat{\mu}, \hat{\sigma}^2)$ of the univariate normal distribution parameters. Start by showing

$$\mathbf{E}[\hat{\mu}] = \mu.$$

(d) Prove that

$$\mathbf{E}[(\hat{\mu} - \mu)^2] = \frac{\sigma^2}{N}.$$

Here, note that μ is not a random variable.

(e) Finally, show that

$$\mathbf{E}[\hat{\sigma}^2] = \frac{N-1}{N}\sigma^2.$$