Questions

[In my presentation of the material in the lecture, I tried to emphasize the concepts more and show lots of pictures and I did not go much into the mathematical calculations. Here in these exercises, I'll push a bit harder on the math.

1. Given a time varying image $I(x, y, t)$, a vision system can estimate partial derivatives with respect to x, y, t in a patch and compute motion constraint equations

   $$\frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} = 0$$

   at each point in the patch. However, if the image intensity $I(x, y, t)$ contains noise then the motion constraint equations will hold only approximately.

   One way to estimate image velocity in the presence of noise is to chose the (v_x, v_y) that minimizes the "cost"

   $$\sum_{(x,y)})(\frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t})^2.$$

   This cost function $cost(v_x, v_y)$ can be thought of as the height of a surface defined on (v_x, v_y) variables. What can you say about the shape of this surface? Does it have a unique minimum?

2. Some computational theories have interpreted the responses of disparity tuned cells at (x, y) in terms of a likelihood function for the disparity at (x, y). Briefly explain how this interpretation might go. Hint: recall the disparity plots that you made for Assignment 2 for the case of the square.

3. Let $\vec{X}$ be a set of independent n random variables which are normally distributed with mean μ and variance σ, so

   $$p(\vec{X} = \vec{x} \mid \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \; e^{-\frac{(x_i - \mu)^2}{2\sigma^2}}$$

   Note that this can be considered as a likelihood function on μ, σ, given an image $\vec{X} = \vec{x}$.

   What is the maximum likelihood estimate of μ for the given observed $\vec{x}$? Hint: Use the exponential to turn the product into a sum.

4. Suppose you have a random variable X that takes value 0 with probably p_0 and 1 with probability $1 - p_0$. Assume $0 < p_0 < 1$.

   Suppose you sample this random variable repeatedly until you get a value 1 and it takes you k attempts, i.e. suppose a 1 first occurs on trial k. The number of samples k before you obtain the first 1 is a random variable: call it K. Write an expression for the likelihood function $p(K = k|p_0)$? That is, what is the likelihood of p_0, given one particular instance $K = k$.

   If you don't know how to calculate the answer to the question, at least make sure you understand what the question is asking.

Solutions

1. If we expand each of the terms in the summation, and then sum them all up, we get an expression of the form
$$A\ v_x^2 + B\ v_y^2 + C\ v_x v_y + D$$
which is a second order polynomial in v_x, v_y, i.e. quadratic. Moreover, inspecting the summation, we see that $A = \sum(\frac{\partial I}{\partial x})^2$ and $B = \sum(\frac{\partial I}{\partial y})^2$ where the sums are over all points in the patch. These sums will be positive and so the sum goes to $+\infty$ when either $v_x \to \pm\infty$ or $v_y \to \pm\infty$.

   Math detail: You might expect that the cost function has a unique minimum. To conclude this, we need to be sure that the noisy gradients $(\frac{\partial I}{\partial x}, (\frac{\partial I}{\partial x})$ are not all parallel to each other. If they are al parallel, then we get a parabolic surface in (v_x, v_y) space where the minimum of the parabola would be motion contraint line that is perpendicular to the unique gradient.

2. A disparity likelihood function is a conditional probability $p(I_l, I_r|d)$ that you observe a stereo image pair given that the disparity is d. This be written in terms of the responses of cells by writing a likelihood $p(R_1, \ldots, R_n|d)$ where $R_1, \ldots, R_n$ are responses of n disparity tuned cells at some image location (x, y) say in the left eye. This is the likelihood of getting these responses when the disparity is d.

   To compute the likelihood, we could approximate the responses as conditionally independent,
$$p(R_1, \ldots, R_n|d) = p(R_1 \mid d), \ldots, p(R_n \mid d)$$
and then we could compute $p(R_i \mid d)$ based on the difference between R_i and the expected value of R_i when the disparity is d, that is, the average of R_i over many examples. Recall from Assignment 2 that you plotted these averages. For example, you could write a likelihood:
$$p(R_i \mid d) = e^{-\frac{(R_i - \mathcal{E}(R_i \mid d))^2}{2\sigma^2}}$$
where σ^2 is the variance of the responses.

3. We can write
$$p(\vec{X} = \vec{x} \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\ e^{-\sum_{i=1}^n (x_i - \mu)^2/\sigma^2}.$$
Take the derivative with respect to μ, set it to 0, and solve for μ. This gives the maximum likelihood for μ is just the sample mean:
$$\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i.$$
Note that its almost never going to happen that $\bar{x} = \mu$ since the sample mean is itself a (continuous valued) random variable.

4. Getting a 1 for the first time on trial k means that you get a 0 on the first $k-1$ trials and then you get a 1 on the kth trial. The trials are independent and so the probability of any particular sequence of 0's and 1's is the product of the probabilities of the individual trials. So, the likelihood function would be $p(K = k|p_0) = p_0^{k-1}(1 - p_0)$. This is unimodal function of p_0 and to find the maximum you take the derivative w.r.t p_0, set it to 0, and solve.