CompSci 527 Final Exam — Sample Solution

1. ***(5 points)*** The last layer of a binary neural net classifier $h : \mathbb{R} \to \{1, 2\}$ computes outptut

$$\mathbf{y} = \begin{bmatrix} 0 \\ \ln 3 \end{bmatrix}$$

in response to some input $\mathbf{x}$. In this expression, $\ln$ denotes the natural logarithm (logarithm in base e). The output $\mathbf{y}$ is passed through a softmax computation to produce vector $\mathbf{p}$. Write numerical values for the entries of $\mathbf{p}$.

Answer:

$$\mathbf{p} = \frac{1}{e^0 + e^{\ln 3}} \begin{bmatrix} e^0 \\ e^{\ln 3} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix} .$$

2. ***(5 points)*** An orthonormal basis for a certain m-dimensional subspace S of $\mathbb{R}^n$, with $m < n$, is given by the columns of the $n \times m$ matrix U.

The orthogonal projection of point $\mathbf{b} \in \mathbb{R}^n$ onto S is a vector $\mathbf{p} = [p_1, \ldots, p_n]^T$ (note that $\mathbf{p}$ still has n coordinates, not m).

Write an expression for $\mathbf{p}$ in terms of $\mathbf{b}$ and U.

Answer:

$$\mathbf{p} = UU^T\mathbf{b} .$$

3. ***(4 points)*** Since the vector $\mathbf{p}$ in the problem above is in S, it can also be written as a linear combination of the m basis vectors in U:

$$\mathbf{p} = \sum_{i=1}^{m} q_i \mathbf{u}_i ,$$

where $\mathbf{u}_i$ is the i-th column of U.

Write a formula, in terms of $\mathbf{b}$ and U, for the vector of coefficients $\mathbf{q} = [q_1, \ldots, q_m]^T$.

Answer: *The entries of* $\mathbf{q}$ *are the inner products of* $\mathbf{b}$ *with the columns of* U*:*

$$\mathbf{q} = U^T\mathbf{b} .$$

4. ***(12 points)*** Steepest descent is used to minimize the function

$$f(\mathbf{x}) = \sin\left(x_1^2 + \frac{\pi}{4}x_2\right) \quad \text{where} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ,$$

starting at point $\mathbf{x}^{(0)} = [0, 0]^T$. Exact line search is used at every iteration. Exact line search is an idealization of line search. It is equivalent to running line search to full convergence, and finding the minimum nearest to the starting point in the search direction. Give the numerical value of the vector $\mathbf{x}^{(1)}$ reached after one line search (that is, after one iteration of steepest descent).

Show your calculations.

Hint: What matters when using line search is the direction of the negative gradient, not its magnitude. Use this fact to keep numbers simple.

Answer: *The gradient of f is*

$$\mathbf{g} = \nabla f = \begin{bmatrix} 2x_1 \\ \frac{\pi}{4} \end{bmatrix} \cos\left(x_1^2 + \frac{\pi}{4}x_2\right)$$

and at $\mathbf{x}^{(0)}$ *the gradient is proportional to the vector*

$$\mathbf{g}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

Therefore, exact line search finds the first minimum of the function

$$\phi(\alpha) = f(\mathbf{x}^{(0)} - \alpha\mathbf{g}_0) = -\sin\left(\frac{\pi}{4}\alpha\right)$$

for $\alpha > 0$.

The derivative of $\phi(\alpha)$ *is*

$$\phi'(\alpha) = -\frac{\pi}{4}\cos\left(\frac{\pi}{4}\alpha\right)$$

and $\phi'(\alpha) = 0$ *for*

$$\alpha = 2 \;+\; 4k \quad \textit{for} \quad k \in \mathbb{Z}\;.$$

Exact line search chooses the first positive solution, $\alpha = 2$*, and we therefore obtain*

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \alpha\mathbf{g}_0 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}\;.$$

5. ***(10 points)*** A tiny CNN layer takes an input $\mathbf{x}$ in $\mathbb{R}^5$ and produces its output $\mathbf{y}$ through a bias of 2, a kernel $[1, 1]$ applied with a "valid" style convolution with a stride of 3, and a ReLU nonlinearity.

   What is the numerical value of the output vector $\mathbf{y}$ of this layer when the input is $\mathbf{x} = [-4, 1, 5, 0, 2]^T$?

   Answer:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \textit{where} \quad y_1 = \max(0, -4+1+2) = 0 \quad \textit{and} \quad y_2 = \max(0, 0+2+2) = 4 \quad \textit{so that} \quad \mathbf{y} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}\;.$$

6. ***(12 points)*** A tiny CNN layer takes an input $\mathbf{x} = [x_1, x_2, x_3, x_4]^T$ and produces its output $\mathbf{y}$ through a kernel $\mathbf{k} = [k_1, k_2]^T$ applied with a "valid" style correlation (not convolution) with a stride of 1, no bias, and no nonlinearity.

   Write the Jacobian matrices

$$J_{\mathbf{x}} = \frac{\partial \mathbf{y}}{\partial \mathbf{x}^T} \quad \text{and} \quad J_{\mathbf{k}} = \frac{\partial \mathbf{y}}{\partial \mathbf{k}^T}$$

   in terms of the entries of $\mathbf{x}$ and $\mathbf{k}$.

   Hint: It is easiest to first spell out the components of $\mathbf{y}$ explicitly in terms of $\mathbf{x}$ and $\mathbf{k}$.

   Answer: *We have*

$$\begin{aligned} y_1 &= k_1x_1 + k_2x_2 \\ y_2 &= k_1x_2 + k_2x_3 \\ y_3 &= k_1x_3 + k_2x_4 \end{aligned}$$

   so that

$$J_{\mathbf{x}} = \begin{bmatrix} k_1 & k_2 & 0 & 0 \\ 0 & k_1 & k_2 & 0 \\ 0 & 0 & k_1 & k_2 \end{bmatrix} \quad \textit{and} \quad J_{\mathbf{k}} = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \\ x_3 & x_4 \end{bmatrix}\;.$$

7. ***(12 points)*** Give numerical expressions for *all* the solutions to the following optimization problem:

$$\arg\min_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \quad \text{where} \quad A = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 0 \end{bmatrix}\;.$$

   Answer: *The matrix has a nontrivial null space, and therefore the problem above admits solutions for which* $\|A\mathbf{x}\|$ *is zero. To be orthogonal to the rows of* A *(and therefore be in the null space of* A*), a vector* $\mathbf{n} \in \mathbb{R}^3$ *must have a first entry equal to 0 (or else it would not be orthogonal to the second row of* A*), and have second and third entry proportional to* $(-1, 2)$ *(so* $\mathbf{n}$ *is perpendicular to the first row of* A*). Thus, the null space of* A *is the line spanned by the vector* $\mathbf{n} = (0, -1, 2)$*. Imposing the unit-norm constraint yields the two solutions*

$$\hat{\mathbf{x}} = \pm\frac{\sqrt{5}}{5}\begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}\;.$$

8. ***(10 points)*** Consider the scalar irradiance on the image sensor of a camera that views a moving scene. "Scalar" here means that there is no color, and image irradiance is the continuous brightness pattern on the sensor, $e(x, y, t)$, a single nonnegative number for every position $(x, y) \in \mathbb{R}^2$ and time $t \in \mathbb{R}$.

   Suppose that, as a result of relative motion between scene and camera, all the points in the image move horizontally to the right at the same speed $u = 2$ pixels per second at all times t, and that the irradiance is everywhere a ramp of constant slope $\frac{\partial e}{\partial x} = 4$ irradiance units per pixel in the horizontal direction. For instance, at some time t_0,

$$e(x, y, t_0) = 4x$$

   (this is a black-and-white image that is dark on the left and gets linearly brighter as you scan the image from left to right). As points in the image move, the irradiance of corresponding points at different times remains constant.

   What is the temporal rate of change $\frac{\partial e}{\partial t}$ of irradiance at any point (x_0, y_0) and time t_0? Also specify the unit of measure for your answer.

   [Note: The answer does not depend on the choice of x_0, y_0, or t_0.]

   Answer: *The optical flow constraint equation can be written for the horizontal component only, since the vertical velocity is zero:*

$$\frac{\partial e}{\partial x} u + \frac{\partial e}{\partial t} = 0$$

   so that

$$\frac{\partial e}{\partial t} = -\frac{\partial e}{\partial x} u = -4 \cdot 2 = -8 \text{ \textit{irradiance units per second.}}$$

9. ***(10 points)*** The Lucas-Kanade algorithm tracks a window centered at point $\mathbf{x}_I$ in image I to its new position in image J by iteratively solving a system of the form

$$A_t \mathbf{d} = \mathbf{b}_t$$

   to find a displacement, and then shifting image J by the solution of that system. Let J_t be the shifted image J just before iteration t.

   Define A_t and $\mathbf{b}_t$. If your definition involves functions not used in this question, define them at least with an English sentence.

   Answer:

$$A_t = \sum_{\mathbf{x}} \nabla J_t(\mathbf{x}) \, [\nabla J_t(\mathbf{x})]^T w(\mathbf{x} - \mathbf{x}_I) \quad \textit{and} \quad \mathbf{b}_t = \sum_{\mathbf{x}} \nabla J_t(\mathbf{x}) \, [I(\mathbf{x}) - J_t(\mathbf{x})] \, w(\mathbf{x} - \mathbf{x}_I) \; .$$

   In these expressions, w is a positive function with a support equal to the window size. It is typically either a constant or an isotropic Gaussian function within its support, and its values add up to one. The summations in the formulas above extend over the entire integer plane.

10. ***(5 points)*** The architecture of FlowNet first computes smaller and smaller feature maps in an *contraction* (or *encoder*) stage, then computes larger and larger maps in an *expansion* (or *decoder*) stage. Name the operations that implement contraction and expansion. No need to mention skip links.

    Answer: *Contraction is achieved with convolution with stride greater than 1. Expansion is achieved with up-convolution with dilution factor greater than 1.*

    [Note: It is not necessary to refer to stride or dilution factor for full credit.]

11. ***(5 points)*** You want to take a portrait in which the subject is crisply in focus and the background is blurred, for a pleasing artistic effect. The subject is fidgety, so you need to take a fast exposure (say 1/100 of a second) to avoid motion blur. Do you achieve better separation between subject and background in dim light or in bright light, assuming that the picture can be properly exposed in either type of lighting by adjusting the lens aperture? You are not allowed to change the exposure time.

    Explain briefly.

    Answer: *In dim light. To achieve good separation it is necessary to open the aperture wide. Since the exposure time is fixed, dim lighting allows opening the aperture wider than in bright light.*

12. ***(10 points)*** Camera 0 and camera 1 are pointed towards the same scene. The center of projection of camera 1 is at position $\mathbf{t} = [2, 0, 0]^T$ measured in the reference system of camera 0. The axes of camera 1 are along the unit vectors

$$\mathbf{i} = \frac{1}{2} \begin{bmatrix} \sqrt{3} \\ 1 \\ 0 \end{bmatrix} \quad , \quad \mathbf{j} = \frac{1}{2} \begin{bmatrix} -1 \\ \sqrt{3} \\ 0 \end{bmatrix} \quad , \quad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

measured in the reference system of camera 0.

Write the numerical value of an essential matrix for the camera pair $(0, 1)$ (in this order).

Answer:

$$E = R[\mathbf{t}]_\times = \begin{bmatrix} \mathbf{i}^T \\ \mathbf{j}^T \\ \mathbf{k}^T \end{bmatrix} \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sqrt{3} & 1 & 0 \\ -1 & \sqrt{3} & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -\sqrt{3} \\ 0 & 2 & 0 \end{bmatrix} .$$

Any nonzero multiple of this matrix is an acceptable answer.