Exam, 2010-2011

Do a total of 22 points (e.g., Q3-Q6, or, Q1 and Q2 and some of Q5, etc.)

Show your work!

Q1: 10 points total; 3 points for part A, 1 point for part B, 2 points each for parts C-E
Q2: 8 points total; 3 points each for parts A and B, 2 points for part C
Q3: 6 points total; 1 point for each of 6 parts
Q4: 6 points total; 1 point for each of 6 parts
Q5: 6 points total; 1 point for each of 6 parts
Q6: 4 points total; 2 points for part A, 1 point each for parts B and C

Q1. A balanced sequence for three kinds of stimuli

Here we design a balanced cyclic sequence for three kinds of stimuli (labeled $\{0,1,2\}$), in which every three-element sequence (except for the sequence $\{0,0,0\}$)occurs exactly once. We do this by extending the finite field $\mathbb{Z}_3$ to make a field of size 27, $GF(3,3)$.

Analogous to $\mathbb{Z}_2$, $\mathbb{Z}_3$ is the field containing $\{0,1,2\}$, with addition and multiplication defined (mod 3). The polynomial $x^3+2x+1=0$ has no solutions in $\mathbb{Z}_3$, so we add a formal quantity ξ to $\mathbb{Z}_3$, and assert that $\xi^3+2\xi+1=0$, and that ξ satisfies the associative, commutative, and distributive laws for addition and multiplication with itself and with the elements of $\mathbb{Z}_3$.

A. Using $\xi^3+2\xi+1=0$, express ξ^r in terms of ξ^0, ξ^1,and ξ^2, for $r=1,...,26$. Note that the sequence of coefficients of ξ^0, considered cyclically, has the property that ever sequence of three labels (except for the sequence $\{0,0,0\}$)occurs exactly once.

B. Consider the sequence of the coefficients of ξ^0 that occur in the expansion of ξ^k, for $k=0,...,25$. Note that the second half of the sequence can be obtained from the first half of the sequence by exchanging 1 and 2. Why must this be true?

C. What is the size of the multiplicative group generated by ξ? Show that for all elements g in this group, $g^{26}=1$.

D. Which of the following maps are automorphisms of the multiplicative group (i.e., are 1-1 maps that preserve multiplication)? $U(g)=g^2$, $X(g)=g^3$, $Y(g)=g^5$

E. Which of the above maps are automorphisms of the field $GF(3,3)$ (i.e., are 1-1 maps that preserve multiplication and addition)?

Q2. Decomposing group representations

Recall that if we have a group G with a representation U_1 in V_1 and another representation U_2 in V_2, we can define a group representation $U_1 \otimes U_2$ on $V_1 \otimes V_2$ by $\left(U_{1,g} \otimes U_{2,g}\right)(v_1 \otimes v_2) = U_{1,g}(v_1) \otimes U_{2,g}(v_2)$. Here we will show that if V_1 and V_2 are the same (i.e., if $V_1 = V_2 = V$) and U_1 and U_2 are the same ($U_1 = U_2 = U$),we can decompose $U \otimes U$ into a symmetric and an antisymmetric component, and determine their characters. We will do this by decomposing $V \otimes V$ into $sym(V \otimes V)$ and $anti(V \otimes V)$, and seeing how each transformation in $U \otimes U$ acts in these components.

A. Recall that a linear transformation A on V acts on a typical element $sym(x \otimes y) = \frac{1}{2}\left(x \otimes y + y \otimes x\right)$ of $sym(V \otimes V)$ by $A\left(sym(x \otimes y)\right) = \frac{1}{2}\left(A(x) \otimes A(y) + A(y) \otimes A(x)\right) = sym(Ax \otimes Ay)$, and similarly for the action of A in $anti(V \otimes V)$. Given that A on V has distinct eigenvalues $\lambda_1, \ldots, \lambda_m$ and eigenvectors $v_1, \ldots, v_m$ (and that V has dimension m), find the eigenvalues and eigenvectors for the action of A in $sym(V \otimes V)$ and $anti(V \otimes V)$ and their traces, denoted $\mathrm{tr}(sym(A \otimes A))$ and $\mathrm{tr}(anti(A \otimes A))$.

B. Our next step is to express $\mathrm{tr}(sym(A \otimes A))$ (i.e., the trace of A acting in $sym(V \otimes V)$), and $\mathrm{tr}(anti(A \otimes A))$ (i.e., the trace of A acting in $anti(V \otimes V)$), in terms of easier quantities. Given the same setup as above, find the trace of $A \otimes A$ (acting in $V \otimes V$) and the trace of A^2 (acting in V), and relate this to $\mathrm{tr}(sym(A \otimes A))$ and $\mathrm{tr}(anti(A \otimes A))$.

C. Finally, recalling that the character is defined by $\chi_L(g) = \mathrm{tr}\left(L_g\right)$, express $\chi_{sym(U \otimes U)}$ and $\chi_{anti(U \otimes U)}$ in terms of χ_U.

Q3: Mutual information: synergy, redundancy, etc.

Consider a stimulus S that is equally likely to have one of several values, $\{0,1,\ldots,N-1\}$, and two neurons that are influenced by it. We only concern ourselves with snapshots, so the responses R_1 and R_2 can be considered to be binary $\{0,1\}$. In this scenario, we can calculate the information conveyed by each neuron alone:

$I_1 = H(S) + H(R_1) - H(S,R_1)$ and $I_2 = H(S) + H(R_2) - H(S,R_2)$,

as well as the information conveyed by the entire "population",

$I_{1+2} = H(S) + H(R_1,R_2) - H(S,R_1,R_2)$.

Note that $H(S,R_1,R_2)$ is the entropy of the table of $4N$ entries, listing the probability $p(S,R_1,R_2)$ of that each value of S is associated with one of the four outputs patterns that R_1 and R_2 can produce (i.e., $\{R_1 = 0, R_2 = 0\}$, $\{R_1 = 1, R_2 = 0\}$, $\{R_1 = 0, R_2 = 1\}$, $\{R_1 = 1, R_2 = 1\}$. This table also determines the joint probabilities of S and each R_i, since $p(S,R_1) = p(S,R_1,R_2 = 0) + p(S,R_1,R_2 = 1)$ and similarly $p(S,R_2) = p(S,R_1 = 0,R_2) + p(S,R_1 = 1,R_2)$.

Find an example of $p(S,R_1,R_2)$ that illustrates each of the behaviors below, or, alternatively, show that the behavior is impossible. It may be work in terms of the stimulus probabilities $p(S)$ and the conditional probabilities $p(R_1,R_2 \mid S)$ (the probability that particular values of R_1 and R_2 occur, given a value of *S*), noting that $p(S,R_1,R_2) = p(R_1,R_2 \mid S)p(S)$.

A. $I_1 > 0, I_2 > 0$, $I_{1+2} = I_1 + I_2$ (independent channels)
B. $I_1 > 0, I_2 > 0$, $I_{1+2} = \max(I_1, I_2)$ (completely redundant channels)
C. $I_1 > 0, I_2 > 0$, $\max(I_1, I_2) < I_{1+2} < I_1 + I_2$ (partially redundant channels)
D. $I_{1+2} > I_1 + I_2$ ("synergistic" channels)
E. $I_{1+2} < \max(I_1, I_2)$ ("occluding" channels)
F. $I_{1+2} < \min(I_1, I_2)$ ("strongly occluding" channels)

Q4. Mutual information with additive noise

Here we will determine the mutual information between a Gaussian stimulus *s* and a response *r* that are related by $r = gs + a$, where *g* is a constant (the "gain), and *a* is an additive noise, uncorrelated with the input. For definiteness, we assume that the stimulus *s* has variance V_S and the noise term is drawn from a Gaussian with variance V_A, and that both have mean 0.

A. For a Gaussian distribution of variance V, $p_V(x) = \frac{1}{\sqrt{2\pi V}} e^{-x^2/2V}$, calculate the differential entropy.

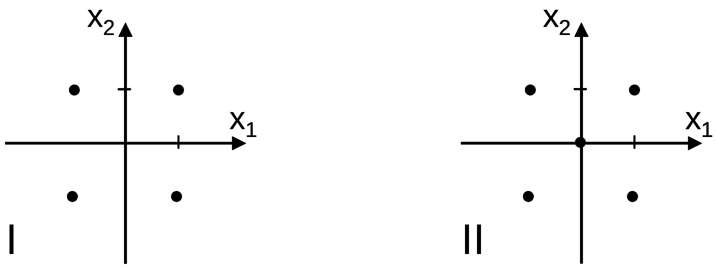
B. How is the output distributed (i.e., what is $p(r)$)? What is its differential entropy?
C. What is the distribution of the output, conditional on a particular value of the input, say, $s = s_0$? I.e., what is $p(r \mid s_0)$? What is its differential entropy?

D. Calculate the mutual information between the input and the output.

E. Calculate the "signal to-noise" ratio, namely, the ratio of the variance of the signal term, *gs*, to the variance of the noise term, *a*. Relate this to the mutual information.

F. Calculate the correlation coefficient between the input and the output,
$C = \frac{\langle (s - \langle s \rangle)(r - \langle r \rangle) \rangle}{\sqrt{\langle (s - \langle s \rangle)^2 \rangle \langle (r - \langle r \rangle)^2 \rangle}}$. Relate this to the mutual information.

Q5. Toy examples of ICA



A. Take four data points, located at $(\pm 1, \pm 1)$, (see diagram I) and consider its projection onto a unit vector $v_\theta = (\cos\theta, \sin\theta)$, to form a distribution of 4 points.

How does the variance of this distribution depend on θ? For which projections is it maximized (equivalently, what direction(s) would be selected by PCA?)

B. As in A, but determine how the kurtosis of the projected distribution depends on θ. Recall, kurtosis is defined by $\kappa = \frac{\left\langle \left(x - \langle x \rangle \right)^4 \right\rangle}{V^2} - 3$, where $V = \left\langle \left(x - \langle x \rangle \right)^2 \right\rangle$ is the variance. For which directions is kurtosis largest?

C. As in A, but determine how the entropy of the projected distribution depends on θ. For which projections is it minimized?

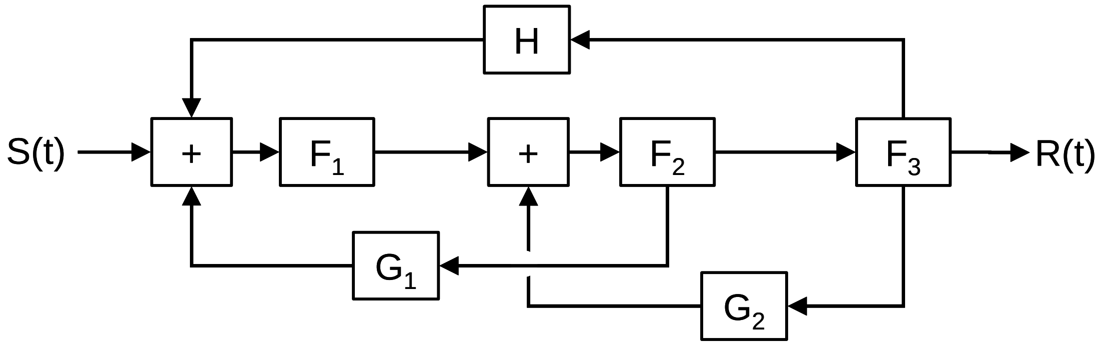
D. Same as in A, but take five data points, located at $(\pm 1, \pm 1)$ and also the orign, (see diagram II) and consider its projection onto a unit vector $v_\theta = (\cos\theta, \sin\theta)$, to form a distribution of 5 points. How does the variance of this distribution depend on θ? For which projections is it maximized (equivalently, what direction(s) would be selected by PCA?)

E. As in B, but determine how the kurtosis of the projected distribution depends on θ. For which directions is kurtosis largest?

F. As in B, but determine how the entropy of the projected distribution depends on θ. For which projections is it minimized?

Q6. Linear systems and feedback

The diagram shows a linear system with input $S(t)$ and output $R(t)$, and the filters F_i, G_i, and H are linear filters with transfer functions $\tilde{F}_i(\omega)$, $\tilde{G}_i(\omega)$, and $\tilde{H}(\omega)$.



A. Find the Fourier transform $\tilde{S}(\omega)$ of $S(t)$ in terms of the Fourier transform $\tilde{R}(\omega)$ of $R(t)$.

B. Write out the relationship between $\tilde{S}(\omega)$ and $\tilde{R}(\omega)$ when (a) the feedback between F_2 and F_1 is absent, (b) the feedback between F_3 and F_2 is absent, or (c) the feedback between F_3 and F_1 is absent.

C. Say that it is known that $F_1 = F_2 = F_3 = F$, and that $G_1 = G_2 = G$. Now consider the four configurations above (the full configuration, and the three configurations of part B, in which one of the three feedback filters has been removed). Do they all produce different outputs?