CS 736 : Medical Image Computing

(Instructor : Suyash P. Awate)

Sample Questions

1. (a) In X-ray computed tomography (CT) imaging, each data point in the transform domain can be understood as the inner product of which two functions (state precisely) ? Describe the semantics underlying both these functions. Name or describe (very briefly) an algorithm for image reconstruction from the acquired data.

   (b) In magnetic resonance imaging (MRI), each data point in the transform domain can be understood as the inner product of which two functions (state precisely) ? Describe the semantics underlying both these functions. Name or describe (very briefly) an algorithm for image reconstruction from the acquired data.

   (c) In diffusion tensor imaging (DTI) of human anatomy, can the diffusion signal $S(g)$, for direction g, be written as a linear function of the tensor D ? What is the functional relationship (give an equation) between $S(g)$ and D ?

   (d) What is a good noise model to use for X-ray CT data in the Radon transform domain ? Justify your model. For this noise model, derive the maximum likelihood (ML) estimate of the noiseless Radon-transform image given N independently repeated observations of the noisy Radon-transform ?

   (e) What is a good noise model to use for X-ray CT data in the spatial domain (NOT Radon-transform domain), assuming reconstructions performed using filtered back projection ? Justify your model. For this noise model, derive the maximum likelihood (ML) estimate of the noiseless CT image image given N noisy CT image reconstructions obtained from repeated imaging experiments ?

   (f) In diffusion MRI, consider a voxel v comprising two fiber tracts, which cross each other at angle θ, where tract-1 is modeled by a diffusion tensor D_1 and tract-2 is modeled by a diffusion tensor D_2. Then, the so-called *multi-tensor* model defines the signal acquired $S(g, D_1, D_2)$, at voxel v, to be a convex combination (superposition) of the two signals $S(g, D_1)$ and $S(g, D_2)$ that would have been acquired for voxels that had only one of the two tracts going through it. $S(g, D_1)$ and $S(g, D_2)$ are described by the single-tensor model that we studied in class. To be able to detect crossings involving tracts that cross at narrow angles θ, how should b values be chosen if (i) the data is noiseless and (ii) the data is noisy ? Describe the trade offs and justify your strategy.

   (g) You have a magnitude image acquired using magnetic resonance imaging (MRI). You want to estimate the noise variance σ^2 from a part of the background region in the image.

   • Derive the maximum likelihood (ML) estimate, in closed form, for the noise level σ assuming the noise model on the magnitude-MRI as (i) Gaussian, (ii) Rician, (iii) Poisson. Which of these estimates will be the most accurate ?

   (h) Suppose you want to register / align two medical images of the same person. Give examples of objective function that would be valid when the images are of (i) identical modalities (e.g., MRI and MRI), (ii) different modalities (e.g., MRI and CT). Justify your answer.

(i) Suppose you want to register / align two medical images from two different persons, but you replace the linear transform with a nonlinear diffeomorphism (i.e., a mapping that is smooth and invertible). Is this a good choice or a bad choice ? Justify your answer. How would the alignment vary as you change the order / strength of smoothness, e.g., from C^1 to C^2 to C^4 to C^8 ?

(j) Suppose you want to align two noiseless shapes (i.e., object boundaries, represented as, say, binary images of very high resolution) using Procrustes alignment, but you replace the similarity transform with a nonlinear diffeomorphism (i.e., a mapping that is smooth and invertible). Is this a good choice or a bad choice ? Justify your answer. Suppose you use this nonlinear transform to compute the mean shape and the covariance of shapes. What would the mean and covariance be ?

2. Consider a maximum-a-posteriori (MAP) image denoising framework that estimates the unknown uncorrupted image x from the observed noisy image y. Consider independent and identically distributed (i.i.d.) zero-mean additive Gaussian noise (known variance σ^2). Consider a Markov random field (MRF) prior (parameters fixed; not to be estimated) on the uncorrupted image that is to be estimated from the noisy data.

   (a) Consider that the chosen MRF model involves non-zero potential functions only for cliques comprising a single pixel. • If the corresponding potential function, at pixel i, is $V_i(x_i) := \beta(x_i)^2$, what is the prior distribution $P(X_i)$ implied on each pixel intensity (give an equation) ? • Specify an algorithm for the MAP denoising, which you think is as computationally efficient as possible.

   (b) Consider that the chosen MRF model involves non-zero potential functions only for cliques involving a single pixel and models all pixel intensities x_i as *non-negative*. • If the corresponding potential function is $V_i(x_i) := \beta(x_i)^2$, what is the prior distribution $P(X_i)$ implied on each pixel intensity (give an equation) ? • Specify an algorithm for the MAP denoising, which you think is as computationally efficient as possible.

   (c) Consider that the chosen MRF model involves non-zero potential functions only for cliques involving a single pixel. • If the corresponding potential function is $V_i(x_i) := \beta|x_i|$, what is the prior distribution $P(X_i)$ implied on each pixel intensity (give an equation) ? • Specify an algorithm for the MAP denoising, which you think is as computationally efficient as possible.

   (d) Consider that the chosen MRF model involves non-zero potential functions only for cliques involving a single pixel and models all pixel intensities x_i as *non-negative*. • If the corresponding potential function is $V_i(x_i) := \beta|x_i|$, what is the prior distribution $P(X_i)$ implied on each pixel intensity (give an equation) ? • Specify an algorithm for the MAP denoising, which you think is as computationally efficient as possible.

   (e) Consider that the chosen MRF model involves non-zero potential functions only for cliques involving two pixels in a standard 4-neighbor (left, right, up, down) neighborhood system. • If the corresponding potential function is $V_i(x_i, x_j) := (x_i - x_j)^2$, what is the *conditional* prior distribution $P(X_i|X_{\sim i})$ implied on each pixel intensity (give an equation) ? • Specify an algorithm for the MAP denoising, which you think is as computationally efficient as possible.

3. Consider a microscopy system that blurs the incoming signal before it reaches the detector. The point spread function associated with the blur is given by the manufacturer. During an imaging experiment, at the detector, a random 20% of the pixels, uniformly spread over the image domain, had their intensity values missing due to some kind of measurement errors.

This means that for such pixels, the pixel intensity data was missing (not to be confused with the value 0). For the remaining pixels, the pixel intensities were corrupted with Poisson noise. You want to restore the missing pixel values as well as reduce the noise and the blur. • Formulate this problem as a MAP optimization problem (give all details for the likelihood and prior models) and specify an algorithm for optimization.

4. The algebraic reconstruction technique (ART) algorithm for X-ray CT reconstruction iteratively updates the estimates of the unknown attenuation-coefficient 2D image as follows. Consider the attenuation-coefficient image matrix to be vectorized (denoted by f) of size $N \times 1$, where N is the number of pixels in the attenuation-coefficient image. At iteration k in ART, the update is defined as $f^{k+1} := f^k + \lambda^k (b_{ij} - \langle a_{ij}, f^k \rangle) a_{ij} / \parallel a_{ij} \parallel^2$, where $b_{ij} \in \mathbb{R}$ is the data at row i and column j in the Radon-transform data matrix and $a_{ij} \in \mathbb{R}^N$ models the integral of the image $f \in \mathbb{R}^N$ with a single ray. • Can this algorithm be understood as a one performing optimization ? If so, give the objective function that such an algorithm aims to optimize. If not, justify why such an objective function cannot be found.

5. Prove that the set of complex waves, i.e., $\{f_n(t) := \exp(int)\}_{n \in \mathbb{I}}$ for integer n, underlying a Fourier series forms a set of orthogonal functions. Find the Fourier-series coefficients in the representation for a real-valued function $g(t) = \sin(\phi) + \sin(t + \theta)$, where θ, ϕ are constants.

6. Assuming perfect data acquisition, i.e., no measurement errors, give the mathematical models of image acquisition underlying (i) X-ray imaging (ii) X-ray Computed Tomography (CT) (iii) Magnetic Resonance Imaging (MRI) (iv) Diffusion Tensor Imaging (DTI). Specifically, describe the physical entity being imaged, describe the kind of data acquired (dimensions and data types), and give the mathematical relationship connecting the former with the latter.

7. Prove, or argue logically, that the backprojection of the Radon transform of an impulse function (in 2D) is the same impulse function.

8. Assume that the matrix-data acquisition in a CT imaging experiment (assume 2D object) gets corrupted such that a random real positive number gets added to a single element of the data matrix. Explain, using theoretical arguments, the effect of this corruption on the inverse Radon transform of the data ? Describe the change in the visual appearance of the image, obtained after the inverse Radon transform, as a result of the corruption in the data.

9. (a) Consider a 3×3 image with (x, y) coordinates $(1, 1)$, $(1, 2)$, $(1, 3)$, $(2, 1)$, $(2, 2)$, $(2, 3)$, $(3, 1)$, $(3, 2)$, and $(3, 3)$, and values $f(x, y)$. Consider a line passing through $(1, 1)$ and $(3, 3)$. Consider performing integration along this line, approximated as a Riemann sum over points $(1, 1), (1.5, 1.5), (2, 2), (2.5, 2.5)$, and $(3, 3)$. Write this line integration as a matrix operation on the vector of image values, i.e., the column vector $[f(x_1, y_1), f(x_1, y_2), f(x_1, y_3), f(x_2, y_1), \cdots, f(x_3, y_3)]'$.

   (b) Describe how you can use this matrix model to write a discretized Radon transform of the image $f(x, y)$. In the context of this matrix-based model, give an algorithm for estimating $f(x, y)$, given the discretized-Radon-transform data, without performing backprojection. Ensure that the proposed algorithm is practically viable.

10. Describe what happens to the aggregate magnetization of Hydrogen protons, within a tissue, when placed in a static magnetic field of strength b_0 oriented along the $+Z$ axis ? Is this process different for different tissues ? What is the aggregation magnetization in steady state ? Is the steady state different for different tissues ?

11. Assume that in a MRI imaging experiment, the matrix data gets corrupted such that a random positive number gets added to the single element of the data matrix. Explain theoretically, the resulting effect on the image of the magnitude values of the inverse-transformed data ?

12. In an MRI experiment, given that the magnetization being imaged is real valued, prove that the acquired data $d(w_1, w_2)$ has the following symmetry: $d(w_1, w_2) = \overline{d(-w_1, -w_2)}$, where $\overline{z}$ denotes the complex conjugate of z and w_1, w_2 denote frequencies.

13. For any non-singular matrix A, is $A'A$ a symmetric positive definite (SPD) matrix ? Prove or disprove. For any non-singular matrix A, is AA' a SPD matrix ? Prove or disprove. For any non-singular square matrix A, is $A + A'$ a SPD matrix ? Prove or disprove.

14. Consider that you acquire corrupted DTI data $S(v_i, g_n)$ for N gradient directions $\{g_n\}_{n=1}^N$ for a set of voxels $\{v_i\}_{i=1}^I$ in a small neighborhood that is known to lie within a single fibrous tissue. You now wish to find a representative diffusion tensor for the selected tissue neighborhood and you consider the following algorithm: Estimate a diffusion tensor $D(v_i)$ at each voxel v_i and compute the element-by-element average of the estimated diffusion tensors (3×3 matrices; 9 elements) and call that a representative diffusion tensor. Is the average also a diffusion tensor ? Prove or disprove. Given the same corrupted data, describe an alternate algorithm to find a representative diffusion tensor ?

15. Prove that the covariance matrix of a vector random variable $X := [X_1, \cdots, X_d]'$ (column vector) is symmetric positive definite (assume that the joint distribution is d dimensional).

16. Consider the algebraic reconstruction technique (ART) for CT image reconstruction, which relies on successive projections of intermediate solutions onto the linear constraint sets represented by hyperplanes (each hyperplane is associated with an integral along a single X ray). • For a simplistic hypothetical scenario involving only 2 such intersecting hyperplanes / constraints, describe how the rate of convergence, of the sequence of intermediate solutions to the optimum, will depend on the angle between the unit vectors normal to the hyperplanes. For instance, what happens when this angle is close to (i) 0 or π (i.e., when the hyperplanes are almost coincident) or (ii) close to $\pm\, \pi/2$ (i.e., when the hyperplanes are almost orthogonal). • In a real-world scenario, how does this behavior dictate the order in which the N hyperplanes / rays should be chosen for the projections to achieve faster convergence? Assume that (i) the distance between parallel rays along the same orientation is much less than the image-pixel width and (ii) the number of orientations over which the X-ray projections are acquired is much larger than 180, uniformly distributed over $(0, \pi)$. Hint: Consider ordering (i) between multiple parallel rays along the same orientation and (ii) between rays with the same distance from the origin (say, object center) but with different orientations.

17. Consider a magnetic resonance imaging (MRI) experiment where the magnetization function $m(x, y)$, over spatial coordinates x, y, is complex valued but with constant phase θ independent of the spatial coordinates, i.e., $m(x, y) := r(x, y)e^{i\theta}$, where $r(x, y) \in \mathbb{R}$. This phase θ is, however, unknown to us. • Assuming an ideal noiseless data acquisition, what is the relationship between the acquired data $d(w_1, w_2)$, at frequencies w_1, w_2, and the acquired data $d(-w_1, -w_2)$? • Assuming a practical noisy acquisition with independent and identically distributed (i.i.d.) zero-mean additive Gaussian noise, how can you use the previously derived relationship to estimate θ (formulate a suitable optimization problem)? • In the practical scenario, how can you use the derived relationship to reduce the noise in the reconstructed image?

18. Consider an MRI acquisition scheme (say, scheme *Cartesian*) where the data is always acquired in "units" where each unit comprises either one row or one column of the 2D grid in the *frequency* domain. Consider an image of size 128×128 pixels in the *spatial* domain; this size also equals the size of the 2D frequency-domain grid. • If you are time constrained to acquire only 32 units of data (i.e., the sum of the number of data columns acquired and the number of data rows acquired is 32), then which of the rows and/or columns will you acquire, and why? Describe the data points acquired on the 2D grid pictorially and logically argue for your proposed scheme.

    Now consider another acquisition scheme (say, scheme *Radial*) that acquires data in units where each unit comprises a radial line (discretized) passing through the origin (i.e., zero frequency) of the 2D frequency domain. Note: the acquired data now lies on a radial grid instead of a Cartesian grid. • If you are time constrained to acquire only 32 units of data (i.e., 32 radial lines), then which radial lines will you acquire, and why? • For such an acquisition, describe clearly how you can mathematically formulate an optimization problem for reconstructing the spatial-domain image represented on a Cartesian grid.

19. Consider a X-ray computed tomography (CT) experiment where N repeated independent measurements $x := \{x_1, x_2, \cdots, x_N\}$ are acquired of a signal X (scalar valued) that is corrupted with Poisson noise. • Derive a formula for an estimator of the uncorrupted signal value, say, y, given the data x, such that the estimator is guaranteed to produce the uncorrupted signal value as $N \to \infty$. Now, consider that some prior information about the uncorrupted value y is known such that $P(Y)$ is itself a Gamma probability mass function with parameters α, β. • Derive another formula for an estimator of the uncorrupted signal value y given the data x as well as the prior $P(Y)$.

20. Consider a D-dimensional multivariate Gaussian random variable $X := AW + b$, where W is a E-dimensional random vector comprising E i.i.d. standard-normal random variables, A is a $D \times E$ constant matrix, and b is a $D \times 1$ constant vector. Consider that we have N independent observations $\{x_1, x_2, \cdots, x_N\}$ drawn from $P(X)$. Consider the method of "principal component analysis" that computes the empirical covariance matrix $\hat{C}$ of the data and performs an eigen analysis of $\hat{C}$ producing eigenvectors $\{v_i\}_{i=1}^{D}$ and eigenvalues $\{\lambda_i\}_{i=1}^{D}$. • Show that, as $N \to \infty$, each v_i is a principal direction of variability of the data and each λ_i is the variance along this direction.

21. Consider a medical image denoising algorithm where the uncorrupted image is modeled using a Markov random field (MRF) X. Consider a Poisson noise model. Consider prior information on the uncorrupted image indicating that the (i) intensities x_i, at pixels i, are spatially piecewise smooth and (ii) each intensity x_i is drawn from a distribution $Q(\cdot)$.

    • Formulate a statistical estimation problem to solve for the uncorrupted image given corrupted data. Clearly specify the prior and the likelihood terms. Also specify the optimization problem.

22. Consider a medical image denoising algorithm where the uncorrupted image is modeled using a Markov random field (MRF). Consider an i.i.d. additive Gaussian noise model. Consider prior information on the uncorrupted image indicating that the intensities are (i) spatially smooth without any discontinuities, (ii) for each pixel, the conditional distribution of the pixel intensity given the intensities in its 8-neighborhood is Gaussian, and (iii) the intensities in the image are drawn from a Gaussian distribution.

    • Formulate a prior model consistent with all the prior information on the uncorrupted image.

    • Formulate a likelihood model.

• formulate an estimation problem to solve for the uncorrupted image given corrupted data. What is the most efficient method to solve the optimization problem ? Justify your answer.

23. A mathematical model of imaging in X-ray computed tomography (CT) can be written as a linear transform $AX = B$, where the X models the image of attenuation coefficients with N pixels, B models the acquired data comprising M measurements, and A is the $M \times N$ matrix modeling X-ray beam integrals through the image X.

    • Assume that $M > N$. In this case, how can we solve for X such that $AX = B$? What conditions are required on A for this to be possible and the solution to be unique ? You cannot include any other information for solving $AX = B$.

    • If we can, in theory, get an X by solving the linear system $AX = B$ in closed form, why cannot this be done in practice ? What can be undesirable about such solutions (mention at least 3 aspects) ?

    • Formulate a Bayesian estimation problem for CT image reconstruction. Use all information available about the practical imaging process and the uncorrupted image x (at least 2 sources of prior information).

    • Suppose, in addition to all the prior information you used in the previous question, you want to learn prior information (through learning on a large dataset of high-quality images) on the typical locations of certain object types in the scene, e.g., bed, patient, air, bones, muscle, fat, etc. Give an algorithm to learn such information from training data ? How will you incorporate this information in a prior MRF model on X ? Formulate precisely using statistical theory.