Homework 5, Stat 541: Due Mon, Oct 8, 2007

Linear Algebra and Linear Inference

Student Name: (replace this with your name)

October 1, 2007

Instructions: Edit this LaTex file with your solutions and generate a PDF file from it. E-mail the PDF to the usual class gmail address.

For some questions one needs to understand basis changes and associated coordinate transformations. To brush up (or to finally really understand what that is), you may want to check the solutions of Homework 2, Problem 11.

1. Interpretation of eigendecompositions: You know that for any real symmetric $p \times p$ matrix $\mathbf{S}$ there exists an orthonormal basis $(\mathbf{u}_j)_{j=1...p}$ of eigenvectors and associated eigenvalues $(\lambda_j)_{j=1...p}$ which can be assumed in descending order w.l.o.g.: $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p$. Simply state what it means for $\mathbf{u}_j$ to be an eigenvector with eigenvalue λ_j.

   Answer:

2. Given $\mathbf{S}$, form the matrix $\mathbf{U} = (\mathbf{u}_1, ..., \mathbf{u}_p)$ and the diagonal matrix Λ that has the eigenvalues in the diagonal. Express what orthonormality of the vectors $\mathbf{u}_p$ means in terms of $\mathbf{U}$. Then express the p eigenvalue conditions with a single matrix equation in terms of, $\mathbf{S}$, $\mathbf{U}$ and Λ.

   Answer:

3. Form the matrices $\mathbf{P}_j = \mathbf{u}_j \mathbf{u}_j^T$ and explain what type of linear map they describe. Then express the p eigenvalue conditions with a single matrix equation in terms of $\mathbf{S}$, $\mathbf{P}_1$,..., $\mathbf{P}_p$ and λ_1,..., λ_p.

   Answer:

4. What is the name of the property of $\mathbf{S}$ called that is equivalent to $\lambda_p \geq 0$? Same for $\lambda_p > 0$?

   Answer:

5. What is the matrix of the linear transformation given by $\mathbf{S}$ in new the new coordinate system after changing the basis to $\mathbf{u}_1$,..., $\mathbf{u}_p$?

   Answer:

6. The following is called the Rayleigh quotient of $\mathbf{S}$:

$$R(\mathbf{u}) \;=\; \frac{\mathbf{u}^T\mathbf{S}\mathbf{u}}{\|\mathbf{u}\|^2}$$

   Find and interpret the stationary equations for $R(\mathbf{u})$: $\partial_{\mathbf{u}} R(\mathbf{u}) = \mathbf{0}$ (∂_u = gradient w.r.t. $\mathbf{u}$).

   Answer:

7. What is the geometric interpretation of the vectors $\cos(t)\mathbf{u}_1 + \sin(t)\mathbf{u}_2$ for any t? If $\lambda_1 = \lambda_2$, try to replace $\mathbf{u}_1$ and $\mathbf{u}_2$ with two such vectors corresponding to two choices t' and t"; what condition on t' and t" is necessary to make this work?

   Answer:

8. What is the eigendecomposition of an orthogonal projection of rank r? How unique is it?

   Answer:

9. What is the matrix of on orthogonal projection after a change to a basis of eigenvectors?

   Answer:

10. Does a non-orthogonal projection have an eigendecomposition as described in Problem 1 above?

    Answer:

11. A fact about traces: show that if $\mathbf{A}$ is $m \times n$ and $\mathbf{B}$ is $n \times m$, then $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$.

    Answer:

12. Use the previous fact to show that traces do not change under basis/coordinate transformation.

    Answer:

13. Using the previous fact, what is the trace of $\mathbf{S}$ in terms of the eigenvalues? Why?

    Answer:

14. Using the previous fact, what is the trace of an orthogonal projection? (Actually, the argument goes through even if $\mathbf{S}$ is not symmetric but has a non-orthogonal basis of the eigenvectors. One can use Problem 2 which is independent of orthogonality. You don't need to do this additional problem, though.)

    Answer:

15. Properties of the diagonal elements of orthogonal projection matrices: Prove $0 \le P_{ii} \le 1$. To this end, write P_{ii} in terms of the matrix $\mathbf{U}$ and the eigenvalues λ_j according to Problem 2 above. Use what you know abou the eigenvalues of projections.

    Answer:

16. If $\mathrm{V}[X]$ is the covariance matrix of a p-dimensional random vector X, express $\mathrm{Var}[\mathbf{a}^T X]$ in terms of $\mathrm{V}[X]$ and $\mathbf{a}$. Hence if you know $\mathrm{V}[X]$, what do you know about the distribution of X in all directions?

    Answer:

17. Two non-negative matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ are said to be ordered, "$\mathbf{S}_1 \ge \mathbf{S}_2$" if $\mathbf{S}_1 - \mathbf{S}_2$ is non-negative definite. If we are given two p-dimensional random vectors X and Y, and if $\mathrm{V}[X] \ge \mathrm{V}[Y]$, what does this mean in view of the previous problem?

    Answer:

18. If in a linear regression problem without intercept the predictor vectors $\mathbf{x}_1$,...,$\mathbf{x}_n$ are i.i.d. samples from the p-dimensional random vector X, is it desirable that $\mathrm{V}[X]$ be "large" or "small" in the sense of the ordering defined in the previous problem? (For clarity: $\mathbf{x}_i$ is the transpose of the i'th row of the predictor matrix.)

    Answer:

19. What does adjustment of a $n \times p$ data matrix $\mathbf{X}$ with regard to $(1, ..., 1)^T \in \mathbb{R}^n$ mean?

    Answer:

20. What does adjustment of a data matrix $\mathbf{X}$ with regard to a categorical variable (such as location in the real estate HW) mean?

Answer:

21. Write an R function that adjusts a matrix `Y` of numeric columns with regard to the variables contained in a dataframe `X`. This is a trivial exercise because you may use the linear models functions `lm()`.

    Answer: