Midterm Announcement

The midterm will be held on Thursday, November 8, 2012. The exam is closed notes and closed book. Bring pen and paper to class. The midterm will cover all of the material discussed in class, the lecture viewgraphs, the reading (in the textbooks and in the Lecture Supplement on Hilbert Space Theory), and the homework assignments up to and including Homework 3. This means that the material on linear vector spaces (up to and including the singular value decomposition, orthogonal projection operators, and the pseudoinverse) can be on the exam.¹

Reading Assignment for Homework 4

Complete your reading of the Lecture Supplement on Hilbert Space Theory. Read the Lecture Supplement on Real Vector Calculus. Be sure to understand the concept of generalized gradient descent algorithms and the use of Lagrange multipliers when performing constrained optimization.

In the textbook by Moon & Stirling, read appendices B, E.1, and E.8 (skip the sections on matrix derivatives for now). Note that this text sometimes uses very confusing notation, such as the use of \( z = y(x) \) on page 896 and entry 6 of table E.1 to denote the fact that \( z \) is a function of \( y \).²

Useful References for Real Vector and Matrix Derivatives

There are several references available which discuss real vector and matrix derivatives. Here are two that I have found very useful, both of which also contain many important vector-matrix identities and matrix properties.

1. *Mathematics for Econometrics*, 3rd Edition, Phoebus J. Dhrymes, Springer, 2000. (Available in paperback.) This text defines (like I do) the vector derivative operator to be a row vector, and is my personal favorite. The third edition has much more advanced material than the second, but the second edition is a model of clear and accessible presentation. (I gave my old copy to my graduate student who uses it all the time.)

2. *Handbook of Matrices*, Helmut Lütkepohl, Wiley, 1996. (Available in paperback.) This is an encyclopedic compendium of matrix facts and identities. It defines the derivative with respect to a vector, \( x \), to be a column vector, but then also defines the vector derivative with respect to the transpose of \( x \), \( x' \), to be a row vector. Thus the latter derivative corresponds to the definition used in class and in the Dhrymes text cited above.

¹The material in this current assignment, Homework 4, will not be covered on the midterm.
²A better notation would be \( z = z(y(x)) \).
Homework Problems

1. Consider the quadratic form

\[ \ell(x) = x^H \Pi x - 2 \text{Re} x^H By + y^H Wy \]

where \( \Pi \) is an \( n \times n \) hermitian, positive-definite matrix, \( \Pi = \Pi^H > 0 \), \( W \) is an \( m \times n \) hermitian, positive-definite matrix and \( B \) is \( n \times m \). Note that with the positive-definite assumptions on \( \Pi \) and \( W \), the quadratic form \( \ell(x) \) is a real-valued function of the complex vector \( x \in \mathbb{C}^n \). Because any real-valued function \( \ell(x) \) of a complex vector \( x \) is non-(complex)-analytic (non-holomorphic), taking the derivative of \( \ell(x) \) with respect to \( x \) is a very touchy issue (see the Lecture Supplement on Complex Vector Derivatives). It is the case, though, that we can minimize the quadratic form shown above by completing the square, thereby side-stepping the need to take derivatives to find its minimum.

a) By completing the square in the manner described in the Lecture Supplement on Hilbert Space Theory, determine the value of \( x \) which minimizes \( \ell(x) \) and the minimum (optimum) value of \( \ell(x) \).

b) Consider the weighted least squares problem

\[ \min_x \| y - Ax \|^2_W \]

where \( x \in \mathbb{C}^n \), \( y \in \mathbb{C}^m \), \( r(A) = n \), and \( W^H = W > 0 \). Solve the weighted least-squares problem by completing the square and show that this solution is the same as obtained via an application of the projection theorem. Give the optimal value of the loss function.

2. a) Write an equivalent table to table E.1 of Moon & Stirling for the case (as defined in class) where the partial derivative with respect to (wrt) a vector is a row operator. (i.e., construct a table of cogradient identities as per Lecture Supplement 2.) b) Prove all the cogradient identities given in your table.

3. a) Rewrite Table E.1 of Moon & Stirling, page 899, using the notation \( \nabla_x \) in place of the partial derivative notation used by Kay. This will place the table in a form consistent with the notation use in the Lecture Supplement 2.\(^3\) If any of the identities seem nonsensical given the mathematical framework that we have developed to date explain why and delete it from the table. b) Prove all the results given in your new table.

4. Restrict Problem 1 above to the real case. Find the optimal solution to the weighted least-squares problem using real vector derivatives. Prove that the optimal solution is a global minimum.

\(^3\)Remember that that Table E.1 is based on the assumption that the space is Cartesian. You can maintain this assumption.
5. Taking \( \hat{\mu} = \mathbf{H}\theta \) in Equation (15.22) of Kay shows that the problem analyzed in Example 15.6, page 521, is equivalent to finding the maximum likelihood estimate of \( \theta \) under the complex gaussian assumption. Find the optimal estimate \( \hat{\theta} \) by minimizing the loss function \( J \) on page 521 using the following various methods. (Note that \( \mathbf{H} \) is assumed to have full column rank.)

(a) The Orthogonality Principle. Find an optimal estimate of \( \hat{\mathbf{x}} \) by projecting onto the subspace spanned by the set of column vectors of \( \mathbf{H} \), \( h_i \), \( i = 1, \ldots, n \). (I.e, ignore the fact that the problem can be treated as a linear inverse problem and instead just view it as a subspace projection problem.)

(b) The geometric approach to solving linear inverse problems. (What we studied most of past few lectures.)

(c) Completing the square as described in the lecture supplement on Hilbert spaces.

(d) One can also solve this problem using complex derivatives as described in the Lecture Supplement on Complex Vector Derivatives. However, we will not consider this possibility.