Exam D0308
Matrix methods
Monday January 16, 2012 Time: kl. 9.00 - 13.00

The final exam consists of: 4 tasks.
Two pages excluding front page
The exam counts 100% of the final grade.
Available aids: pen and paper

Teacher: PhD David Di Ruscio
Systems and Control Engineering
Department of technology
Telemark University College
N-3914 Porsgrunn

Task 1 (25%): The four fundamental subspaces

Assume given a matrix, $A \in \mathbb{R}^{m \times n}$, and a linear equation, $Ax = b$, where vectors, x, and, b, have compatible dimensions.

a) What is the dimensions of the vectors x and b ?
Solution: Dimensions $x \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$.

b) What is meant with the rank, r, of the matrix A ?.
Solution: The rank, r, of a matrix $A \in \mathbb{R}^{m \times n}$ satisfy $0 \leq r \leq \min(m, n)$. If $m \geq n$ then the rank r is the number of linearly independent columns. In general the rank r is the number of linearly independent rows or columns of the matrix A. Furthermore the rank, r of matrix A is equal to the number of pivots (p. 144) and equal to the number of non-zero singular values of matrix A.

c) Define each of the four fundamental subspaces.
Solution: Se Ch. 3.6. The row space $C(A^T)$. The column space $C(A)$. The nullspace $N(A)$ The left nullspace $N(A^T)$.

d) Specify the dimension of each of the four fundamental subspaces.
Solution: Se Ch. 3.6. The dimension of the row space is the rank, r. The dimension of the column space is the rank, r. The nullspace has dimension $n - r$. The left nullspace has dimension $m - r$.

e) Give a general requirement for the linear equation, $Ax = b$, to have a unique solution x.
Solution: p. 159. When A is square and invertible, then $r = m = n$ and $Ax = b$ has one solution.

Task 2 (25%): Orthogonality

Assume given a matrix, $A \in \mathbb{R}^{m \times n}$.

a) Discuss the concept of orthogonality of the four fundamental subspaces.
Solution:
Se Ch. 4.1.

- The row space $C(A^T)$ is perpendicular to the nullspace $N(A)$. The nullspace $N(A)$ and the row space $C(A^T)$ are orthogonal subspaces of $\mathbb{R}^n$.

- The column space $C(A)$ is perpendicular to the nullspace of A^T (the left nullspace $N(A^T)$). The column space $C(A)$ and the left nullspace $N(A^T)$ are orthogonal subspaces in $\mathbb{R}^m$.

b) Discuss the concept of projections in connection with the linear equation, $b = Ax + e$, where, e, is the error vector.
Hint: answer should include: projection matrix P, the solution $\hat{x}$, the projection of b onto the subspace of A and the error $b - A\hat{x}$.
Solution:
Se p. 210.

- The least squares solution is, $\hat{x} = (A^T A)^{-1} A^T b$.
- The projection of b onto the column space of A is $\hat{b} = A\hat{x} = p = Pb$ where the projection matrix then is $P = A(A^T A)^{-1} A^T$.
- The error $\hat{e} = b - \hat{b} = b - A\hat{x} = b - Pb = (I - P)b$ is perpendicular to the column space of A. In other words: The error $b - \hat{b}$ is the projection of b onto the orthogonal complement of A.

c) Give a short description of the QR decomposition of the matrix, A.
Solution:
Se pp. 235-237.

- A matrix A may be decomposed (factorized) into $A = QR$ where Q is an orthogonal matrix such that $Q^T Q = I$ and the columns in Q are orthonormal vectors.
- The matrix $R = Q^T A$ is upper triangular.
- The QR decomposition may be effectively computed using the Gram-Schmidt method, p. 237.

d) Consider a linear equation, $b = Ax + e$, where, $A \in \mathbb{R}^{m \times n}$, and $m > n$.

Show how the QR decomposition of the concatenated matrix

$$\begin{bmatrix} A & b \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} R_{11} & R_{21} \\ 0 & R_{22} \end{bmatrix}, \tag{1}$$

can be used to find the least squares solution, $\hat{x}$, to x?
Solution:
We have from Eq. (1) that $A = Q_1 R_{11}$ and $b = Q_1 R_{21} + Q_2 R_{22}$. Using this in $b = Ax + e$ we have $Q_1 R_{21} + Q_2 R_{22} = Q_1 R_{11} x + e$.

Multiplying this last eq. with Q_1^T gives the equation $R_{21} = R_{11} x$ because $Q_1^T Q_1 = I$, $Q_1^T Q_2 = 0$ and we assume $Q_1^T e = 0$.

Hence, the least squares solution may be computed as $\hat{x} = R_{11}^{-1} R_{21}$.

Notice that this is a robust and effective algorithm, in particular when the number of rows m is much larger than the number of columns n.

Se also p. 236 (middle part) for similar details regarding the QR and least squares.

Task 3 (25%): Singular Value Decomposition (SVD), norms and linear regression

a) Discuss the singular value decomposition of a matrix, $A \in \mathbb{R}^{m \times n}$.
Solution:
Se Ch. 6.7 p. 363.

- The SVD of a matrix A is $A = U\Sigma V^T$ where U is a singular vector matrix such that $U^TU = I$, V is a singular vector matrix such that $V^TV = I$ and Σ is diagonal matrix with the $p = \min(m, n)$ singular values $\Sigma_i \geq 0$ on the diagonal.
- Consider the matrix AA^T which is $AA^T = U\Sigma^2U^T$. Hence, U is the eigenvector matrix of the symmetric matrix AA^T and the singular values is the square root of the eigenvalues of the symmetric matrix AA^T.
- Similarly from the matrix $A^TA = V\Sigma^2V^T$ we se that V is the eigenvector matrix for the symmetric matrix A^TA.
- In the above discussion the eigenvalue matrix Σ^2 is actually the diagonal matrices $\Sigma\Sigma^T$ and $\Sigma^T\Sigma$, respectively.

b)

- Explain what is meant with the length (or norm), $\|E\|$, of a vector, $E \in \mathbb{R}^m$.
  Solution:
  Se p. 12. The length or norm of a vector is $\|E\| = \sqrt{E^TE}$.
- Explain what is meant with the Frobenius norm, $\|E\|_F$, of a matrix, $E \in \mathbb{R}^{m \times n}$.
  Solution:
  Se p. 475 Ch. 9.7 where we find the only place where the Frobenius norm is described.

  The Frobenius norm is equal to the square root of the sum of the square of all the elements of A, i.e. $\|E\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2} = \sqrt{\text{trace}(A^TA)}$.

c) Consider a linear equation, $Y = XB + E$, where, $X \in \mathbb{R}^{N\times n}$, and $N > n$ and, $r = \text{rank}(X) < n$, and where we assume that Y is a vector.

Show how the Singular Value Decomposition (SVD) of the matrix, X, can be used to find the Principal Component regression (PCR) estimate, $\hat{B}_{\text{PCR}}$ of B.
Hint: The solution should minimize the squared length (or Frobenius norm), $\|E\|^2 = \|E\|_F^2$ when the error $E = Y - XB$ is a vector, and where the estimated error is $\hat{E} = Y - X\hat{B}_{\text{PCR}}$.
Solution:
Here we take the SVD of matrix X, i.e. $X = U\Sigma V^T \approx U_1\Sigma_1 V_1^T$ and solve $Y = U_1\Sigma_1 V_1 B$ for B which gives $\hat{B}_{\text{PCR}} = V_1\Sigma_1^{-1}U_1^T Y$.

Task 4 (25%): Eigenvalues and the QR method

Assume given a square matrix, $A \in \mathbb{R}^{n\times n}$.

a) Discuss and define the eigenvalue decomposition of the matrix A.
Tips: answer should include eigenvalues, eigenvectors, the eigenvalue matrix, Λ, and the eigenvector matrix, S.
Solution:

- Se p. 298 Ch. 6.2.
A square matrix $A \in \mathbb{R}^{n\times n}$ with distinct eigenvalues $\lambda_i \; \forall \; i = 1, \ldots, n$ (in general, with linearly independent eigenvectors $x_i \; \forall \; i = 1, \ldots, n$) may be decomposed as $A = S\Lambda S^{-1}$ where S is an eigenvector matrix with the eigenvectors x_i as columns, and Λ is a diagonal eigenvalue matrix with the eigenvalues λ_i on the diagonal.
- Se p. 287.
The eigenvalues λ_i may be computed by solving the characteristic equation $\det(A - \lambda I) = 0$.
Remark: Solving the characteristic equation is equivalent of solving an nth order polynomial for the n eigenvalues $\lambda_i \; \forall \; i = 1, \ldots, n$. This is an inaccurate and slow (a terrible) method of calculating the eigenvalues (p. 487).
- Se p. 288.
Furthermore, for each eigenvalue λ_i, the eigenvectors x_i may be defined through the linear equation $Ax_i = \lambda_i x_i$.

b)

- What is the eigenvalues of the transpose A^T of the matrix A ?
  Solution:
  Se p. 295.
  From $A = S\Lambda S^{-1}$ we find $A^T = S^{-T}\Lambda S^T$ and hence the eigenvalues of A^T is equal to the eigenvalues of the A matrix.
- Define the trace, trace(A), as a function of the n eigenvalues of matrix A?
  Definition: The sum of the entries of the main diagonal is called the trace of A, i.e. trace(A).
  Solution:
  Se p. 289.
  We have trace$(A) = \sum_{i=1}^{n} \lambda_i$, i.e., the trace is equal to the sum of the eigenvalues.
- Define the determinant, det(A), as a function of the n eigenvalues of the matrix A?
  Solution:
  Se p. 295.
  We have det$(A) = \prod_{i=1}^{n} \lambda_i$, i.e. the determinant of A is equal to the product of the n eigenvalues.

c) Discuss the QR method for calculating the eigenvalues. **Solution:**
Se p. 287.
The QR method for calculating the n eigenvalues $\lambda_i \ \forall \ i = 1, \ldots, n$ is one of the most amazing algorithms in the history of linear algebra.

The algorithm is briefly as follows:

- Factor A into $A = QR$ where Q has orthonormal columns and R is upper triangular.
- Reverse Q and R and form a new matrix $A_1 = RQ$. Notice that A and A_1 has the same eigenvalues because they are similar matrices and $A_1 = Q^{-1}AQ$, because $R = Q^{-1}A$.
- Factor A_1 into the QR decomposition $A_1 = Q_1R_1$
- Reverse the matrices and form the new matrix $A_2 = R_1Q_1$. Notice that A_1 and A_2 has the same eigenvalues because they are similar matrices and $A_2 = Q_1^{-1}A_1Q_1$.
- Continuing this iterative process for a number i of iterations, until convergence. Then the eigenvalues are located on the diagonal of A_i if the eigenvalues are real. If complex eigenvalues they may be located as 2×2 blocks on the diagonal of A_i. Remark that A_i usually is real and upper block triangular and that we may take $T = A_i$ and that we at this stage have computed the amazing block real Schur decomposition $A = QTQ^T$.

A MATLAB function implementation is given below in order to illustrate the QR method.

```
function [T,Qt]=qr_it2(A);
% QR_IT2 Given an (n x n) matrix A.
% This function are using QR iterations in order to calculate the
% amazing block real Scur decomposition A=Q*T*Q'.
% on output.
% [T,Q]=qr_it2(A)
% ON INPUT
% A - An (n x n) matrix with real eigenvalues
% ON OUTPUT
% T - An (n x n) diagonal or upper triangular matrix similar to A,
% i.e. with the n-eigenvalues of A on the diagonal.
% Q - An orthogonal matrix such that A=Q*T*Q'

% Written: November 10 2011 by David Di Ruscio

n=size(A,1);
Qt=eye(n);

max_it=100;
for i=1:max_it
    [Q,R]=qr(A);
    A=R*Q;
    Qt=Qt*Q;
end
T=A;
```