E3101: A study guide and review, Version 1.5

Marc Spiegelman

December 7, 2004

Here is a list of subjects that *I* think we've covered in class (your mileage may vary). If you understand and can do the basic problems in this guide you should be in very good shape. This guide is probably over-thorough. The test itself will have about 6-7 questions covering the whole course but emphasizing the material after the midterm (which implicitly includes all of the previous material). I'll try to avoid anything *overly* tricky. This initial version may have some mistakes…Keep an eye on the version number.

1 Chapter 2: Solution of $A\mathbf{x} = \mathbf{b}$ for square A

Elimination and the LU decomposition For any square matrix A know how to factor A into $PA = LU$ where P is a permutation matrix, L is a lower triangular matrix of multipliers l_{ij} with 1's on the diagonal and U is a an upper triangular matrix with Pivots on the diagonal.

An example of a matrix that needs a row exchange is

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 3 \\ 1 & 0 & 0 \end{bmatrix} \tag{1}$$

as the second row will definitely cause a zero in the 2,2 position after elimination. There is no *unique* permutation required, however and you can always get creative as long as it works (i.e. doesn't generate another zero in a pivot position). For example one possible permutation is to move row 3 to row 1, row 1 to row 2 and row 2 to row 3 i.e.

$$PA = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 3 \\ 2 & 4 & 3 \end{bmatrix} \tag{2}$$

which is now readily factored to

$$PA = LU = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & -3 \end{bmatrix} \tag{3}$$

Solution of Square $A\mathbf{x} = \mathbf{b}$ Also be able to solve the basic square problem $A\mathbf{x} = \mathbf{b}$ by elimination by forming the augmented matrix $[A \;\; \mathbf{b}]$, eliminating to $[U \;\; \mathbf{c}]$ then solving by back-substitution.

Find A^{-1} by Gauss Jordan Elimination You may need this to do eigenvalue problems. First form the augmented matrix $[A \;\; I]$ where I is the identity matrix, then proceed to eliminate the entire matrix downwards, then upwards, then divide by the pivots to end up with $[I \;\; A^{-1}]$.

A^{-1} for a 2 by 2 matrix The formula for the inverse of a general 2×2 matrix is useful to remember, particularly for eigenproblems.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{4}$$

Product rules for inverse (and transpose) If matrices A and B are *both* invertible, then $(AB)^{-1} = B^{-1}A^{-1}$ (A and B must both be square to be invertible). Product rule for transpose looks the same although A and B do not have to be square or invertible, i.e $(AB)^T = B^T A^T$.

2 Chapter 3: Solutions of general $m \times n$ $A\mathbf{x} = \mathbf{b}$ and the four subspaces

General solution of $A\mathbf{x} = \mathbf{b}$ For a general $m \times n$ matrix, know how to find all solutions of the linear system $A\mathbf{x} = \mathbf{b}$. The general solution is to form the augmented matrix $[A \;\; \mathbf{b}]$ and use Gauss-Jordan elimination to reduce it to *Row Reduced Echelon Form* $[R \;\; \mathbf{d}]$. Then

1. Identify the pivot columns and the free columns
2. Determine the rank of the matrix r (number of pivot columns)
3. if $r < n$ (there are free columns), find the null space $N(A)$ by finding the combination of *pivot* columns that cancel each free column.
4. Find a particular solution $\mathbf{x}_p$ as the combination of pivot columns of R that combine to form the right hand side d. (Note: the particular solution $\mathbf{x}_p$ is not necessarily entirely in the Row space of A (i.e. see SVD).
5. If $\mathbf{d}$ is not in the column space of R. The problem has **NO** solution.

6. If A has a non-zero null space then there are an **infinite** number of solutions $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_N$ where $\mathbf{x}_N = N\mathbf{c}$ is in the Null space spanned by the columns of the matrix N.

an Example: if the row reduced echelon form of A is

$$R = \begin{bmatrix} 1 & -3 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{with} \quad \mathbf{d} = \begin{bmatrix} 2 \\ -3 \\ a \end{bmatrix} \tag{5}$$

Then the matrix has *rank* $r = 2$. Columns 1 and 3 are pivot columns, 2 and 4 are free columns. The dimension of the null space is $n - 4 = 2$ with a basis forming the columns of

$$N = \begin{bmatrix} 3 & -2 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \tag{6}$$

if $a \neq 0$ this problem has **no** solution. If $a = 0$ then it has the general solution

$$\mathbf{x} = \begin{bmatrix} 2 \\ 0 \\ -3 \\ 0 \end{bmatrix} + N\mathbf{c} \tag{7}$$

where $\mathbf{c} = (c_1, c_2)$ is any real vector in $\mathbf{R}^2$.

Independence, Basis and dimension Understand how to determine if a set of vectors are *linearly independent* (i.e. if the vectors form columns of a matrix A, then they are linearly independent if the only solution to $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$). A *Basis* for a vector space is a *linearly independent* set of vectors that *span* the space (i.e. any linear combination of the basis vectors fill the entire space). The *dimension* of a subspace is the number of basis vectors required to describe it.

- Any space can have an infinite number of bases (but they all have the same dimension)
- Given a vector space of dimension n with a particular basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, however, any vector $\mathbf{x}$ in the space can be *uniquely* decomposed into a linear combination of the basis vectors.

The 4 fundamental subspaces of A Every $m \times n$ matrix A has associated with it four fundamental subspaces

1. The Column Space $C(A)$ which is the linear combination of the columns of A.

2. The Row Space $C(A^T)$ which is the linear combination of rows of A (i.e. columns of A^T.)
3. The Null Space $N(A)$. The vector space of solutions $\mathbf{x}$ to $A\mathbf{x} = \mathbf{0}$
4. The Left Null Space $N(A^T)$. The vector space of solutions $\mathbf{x}$ to $A^T\mathbf{x} = \mathbf{0}$

Some important factoids about the 4 subspaces

- The **row** space and the **null** space are orthogonal to each other and together form a basis for $\mathbf{R}^n$
- The **column** space and the **left null** space are orthogonal to each other and together form a basis for $\mathbf{R}^m$
- The Column space and the row space both have dimension r (where r is the rank of A). $\dim(N(A)) = n - r$, $\dim(N(A^T)) = m - r$.
- $A\mathbf{x}$ takes vectors in the row space and maps them to the column space and takes vector in the Null space and maps them to $\mathbf{0}$.
- Only for invertible matrices does $A^{-1}\mathbf{b}$ map the column space back to the row space (But see the SVD).

Important Know how to find the dimensions and bases of the 4 subspaces from any matrix A and its row reduced echelon form R (for the left null space, the easiest way to find it is to find the null space of A^T). Also know how to find orthogonal bases for the 4 subspaces from the SVD.

3 Chapter 4: Projections, Least Squares Problems, Gram-Schmidt Orthogonalization and the QR decomposition

You should know how to do the following things…

Projection onto a line If a vector $\mathbf{a}$ forms the basis for a line ($c\mathbf{a}$), then the projection of another vector $\mathbf{b}$ onto that line is given by

$$\mathbf{p} = \mathbf{a}\hat{x} \tag{8}$$

where

$$\hat{x} = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} \tag{9}$$

We can also write the projection as

$$\mathbf{p} = \left(\frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}\right)\mathbf{b} = P\mathbf{b} \tag{10}$$

where P is the projection matrix that only depends on $\mathbf{a}$. Note the projection matrix is singular and just picks out the component of $\mathbf{b}$ that is parallel to $\mathbf{a}$. P project the part of $\mathbf{b}$ that is perpendicular to $\mathbf{a}$ to zero (i.e. that part is in the Null space of P)

Projection onto a subspace A line is just a 1-D subspace spanned by a single vector. We can extend the idea of projection onto a n dimensional subspace described by a larger number of basis vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. If we form a matrix A whose columns are the basis vectors, then the projection of a vector $\mathbf{b}$ onto the column space of A is the point in A that is closest to $\mathbf{b}$ (i.e. minimizes the length of the error vector $\mathbf{e} = \mathbf{b} - \mathbf{p}$). The projection is

$$\mathbf{p} = A\hat{\mathbf{x}} \tag{11}$$

where $\hat{\mathbf{x}}$ is the solution of

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b} \tag{12}$$

This has a *unique* solution if A is full column rank (i.e. $A^T A$ is invertible). If A has a null-space it is not unique but there is always a shortest solution given by the pseudo-inverse (see the last section). Equations (11)–(12) can also be written in terms of a projection matrix

$$\mathbf{p} = P\mathbf{b} = [A(A^T A)^{-1}A^T]\mathbf{b} \tag{13}$$

(if A is full column rank, i.e. $A^T A$ is invertible. If A is not full column rank $P = AA^+$ where A^+ is the pseudo-inverse). Note: **Most projection matrices are singular!** i.e. they will have a null space $P\mathbf{x} = 0$ whenever $\mathbf{x}$ is orthogonal to the column space of A (or whenever there is a left null space $A^T = 0$.) However, if A is square and invertible then $P = I$.

Application of projections #1: Least Squares solutions We can use the idea of projections to solve *least squares problems* where $A\mathbf{x} = \mathbf{b}$ doesn't have a solution because **b** does not lie in the column space of **A**. The best solution $\hat{\mathbf{x}}$, which minimizes the error $\mathbf{e} = A\hat{\mathbf{x}} - b$ is just the solution of

$$A^T A\hat{\mathbf{x}} = A^T b \tag{14}$$

Example...fitting of a straight line to data Suppose you wanted to fit a straight line

$$y = c_1 + c_2 x \tag{15}$$

through a set of 4 (x, y) points $(-3, 1)$, $(-1, 1)$, $(0, 0)$, and $(2, 4)$. Each point i is trying to satisfy an equation $y_i = c_1 + c_2 x_i$ or all 4 points form the system of equations

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{16}$$

or

$$\mathbf{y} = A\mathbf{c} \tag{17}$$

For the specific points here $\mathbf{y} = (1, 1, 0, 4)$ and

$$A = \begin{bmatrix} 1 & -3 \\ 1 & -1 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} \tag{18}$$

To solve for the best fitting parameters c_1, c_2. Just solve the least squares problem $A^T A\mathbf{c} = A^T\mathbf{y}$ or

$$\begin{bmatrix} 4 & -2 \\ -2 & 14 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix} \tag{19}$$

by elimination (and then show that the error is in the left null space of A). Also know how to solve in general for best fit quadratic (or general polynomials) of form

$$y = c_1 + c_2 x + c_3 x^2 \tag{20}$$

Most generally, if you want to find the least-squares fit of function as a linear combination of more general functions (e.g. $\sin$ and $\cos$) such as

$$y = c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) \tag{21}$$

through points (x_1, y_1), (x_2, y_2),…,(x_n, y_n) you need to find the least-squares solution of $A\mathbf{c} = \mathbf{y}$ that looks like

$$\begin{bmatrix} f_1(x_1) & f_2(x_1) & f_3(x_1) \\ f_1(x_2) & f_2(x_2) & f_3(x_2) \\ \vdots & \vdots & \vdots \\ f_1(x_n) & f_2(x_n) & f_3(x_n) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{22}$$

Application of Projections #2: Gram-Schmidt orthogonalization Also know how to take an arbitrary basis for a vector space and find an orthonormal basis for it using Gram-Schmidt orthogonalization. I.e. given a set of basis vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ find an orthonormal set of vectors $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ that span the same space and have the property that

$$\mathbf{q}_i^T \mathbf{q}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{23}$$

i.e. if the **a**'s form the column space of a matrix A, find a matrix Q whose columns are the **q**'s and is orthonormal such that $Q^T Q = I$.

The basic algorithm is *Gram-Schmidt Orthogonalization*, which is a process of sequential projections that effectively straighten out the vectors one-by-one. In short form, the first 3 steps are

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{||\mathbf{a}_1||} \tag{24}$$

$$\mathbf{b}_2 = \mathbf{a}_2 - (\mathbf{q}_1^T \mathbf{a}_2)\mathbf{q}_1 \qquad \mathbf{q}_2 = \frac{\mathbf{b}_2}{||\mathbf{b}_2||} \tag{25}$$

$$\mathbf{b}_3 = \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3)\mathbf{q}_1 - (\mathbf{q}_2^T \mathbf{a}_3)\mathbf{q}_2 \qquad \mathbf{q}_3 = \frac{\mathbf{b}_3}{||\mathbf{b}_3||} \tag{26}$$

Application #3: the QR decomposition and least squares solutions Given Gram-Schmidt to calculate Q from A you can find the QR factorization $A = QR$ simply by calculating the upper triangular matrix $R = Q^T A$ (because $Q^T Q = I$). Then, the least squares problem $A^T A\hat{\mathbf{x}} = A^T \mathbf{b}$ becomes $R\hat{\mathbf{x}} = Q^T \mathbf{b}$ which can be solved quickly for $\hat{\mathbf{x}}$ by back substitution.

Some Additional Q matrix Properties

- $Q^T Q = I$ for all $m \times n$ Q matrices
- **For square matrices only** $Q^T = Q^{-1}$ and therefore $QQ^T = I$
- For non-square matrices, $P = QQ^T$ is the projection matrix onto the column space of Q.

- If $A = QR$ then QQ^T is also the projection matrix onto the columns space of A.
- Multiplication by Q matrices doesn't change the length of vectors, i.e. $||Q\mathbf{x}|| = ||\mathbf{x}||$.
- Transformations by Q matrices preserve dot products. i.e. if $\mathbf{x}' = Q\mathbf{x}$ and $\mathbf{y}' = Q\mathbf{y}$ then $(\mathbf{x}')^T\mathbf{y}' = \mathbf{x}^T Q^T Q\mathbf{y} = \mathbf{x}^T\mathbf{y}$
- Rotation, reflection and permutation matrices are all examples of Q matrices.

4 Chapter 5: Determinants

The determinant of a *Square* matrix A (written $\det(A)$ or $|A|$) is a single number formed by combinations of all the elements of A. The determinant has the following useful formulas and properties

- the determinant of a 2×2 matrix
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{27}$$
is $|A| = ad - bc$
- The determinant of an $n \times n$ matrix can be determined by the *Cofactor formula* along any row or column (see Strang 222–224).
- The determinant of an $n \times n$ matrix can also be found by first eliminating A to its upper triangular form U, the taking the **product of the pivots**. This is probably the fastest way for larger matrices.
- **Important!** the determinant of a singular matrix is 0. (i.e. if $A\mathbf{x} = \mathbf{0}$ for some $\mathbf{x} \neq \mathbf{0}$ then $\det(A) = 0$
- the determinant of the identity matrix is 1 ($\det(I) = 1$)
- the determinant changes sign with a single row (or column) exchange
- the determinant is a linear function of each row separately (see Strang 209)
- If 2 rows (or columns) of a matrix A are the same, $\det(A) = 0$
- Elimination (without row exchanges) does not change the determinant
- **Product Rule** (very **important**) $\det(AB) = \det(A)\det(B)$
- if $PA = LU$ then $\det(A) = \det(P)\det(U)$ where $\det(P) = \pm 1$ depending on whether the permutation matrix has an even or odd number of row exchanges and $\det(U)$ is the product of the pivots!

- $\det A^{-1} = 1/\det(A)$
- $\det A^T = \det A$
- $\det(cA) = c^n \det A$
- $\det A =$ the product of the pivots (= the product of the eigenvalues of A).
- **Useful!** The determinant of a triangular matrix is just the product of the diagonal terms.
- The main application of the determinant is finding Eigenvalues by solving $\det(A - \lambda I) = 0$ for n eigenvalues λ.

5 Chapter 6: Eigenproblems and applications

Eigenvalue, Eigenvector problems are solutions of the equations

$$A\mathbf{x} = \lambda\mathbf{x} \tag{28}$$

where A is a **square** $n \times n$ matrix. Equation (28) is slightly deceptive as it is actually an equation for n eigenvalues and eigenvectors.

Basic recipe for finding eigenvalues and eigenvectors

1. Find all **eigenvalues** by finding the n roots of the characteristic equation given by $\det(A - \lambda I) = 0$ (i.e. find the n λ's that make $A - \lambda I$ singular.)
2. Find each **eigenvector**. For each eigenvalue λ_i find the special solutions (i.e. Null space) of the matrix $A - \lambda_i I$

An example: find the eigenvalues and eigenvectors of the 2×2 singular matrix

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \tag{29}$$

1. $\det(A - \lambda I) = 0$ implies that $(1-\lambda)(6-\lambda) - 6 = \lambda^2 - 7\lambda + (6-6) = 0$ or $\lambda_1 = 0$ and $\lambda_2 = 7$ (the order of the eigenvalues is arbitrary). Clearly A is a singular matrix.
2. The first eigenvector is the special solution of the null space of $A - 0I$ (i.e. find the Null space of A), Either by inspection or by reducing to row reduced echelon form, we find $\mathbf{x}_1 = (-2, 1)$.

3. The second eigenvector is in the null space of

$$A - 7I = \begin{bmatrix} -6 & 2 \\ 3 & -1 \end{bmatrix} \tag{30}$$

which is (hopefully clearly) $\mathbf{x}_2 = (1, 3)$ (i.e. the first column plus three times the second column of $A - 7I$ is (0,0).

Useful checks for eigenvalues

1. the **sum** of the eigenvalues is equal to the **trace** of A (the sum of the diagonal terms of A). (i.e. $\mathrm{Tr}(A) = \sum_{i=1}^{n} \lambda_i$)
2. The **product** of the eigenvalues is equal to the **determinant** of A (i.e. $\det(A) = \Pi_{i=1}^{n} \lambda_i$

Eigenvectors and Eigenvalues of a general 2×2 matrix This shows up enough time to be worth remembering now and again (but always know how to derive it).

Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the eigenvalues are given by

$$\lambda = \frac{\mathrm{Tr}(A) \pm \sqrt{\mathrm{Tr}(A)^2 - 4\det(A)}}{2}$$

which are just the roots of

$$\lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = 0$$

Given λ_1 and λ_2, the two eigenvectors can be found with either

$$\mathbf{x}_1 = \begin{bmatrix} b \\ \lambda_1 - a \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} b \\ \lambda_2 - a \end{bmatrix}$$

or

$$\mathbf{x}_1 = \begin{bmatrix} \lambda_1 - d \\ c \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} \lambda_2 - d \\ c \end{bmatrix}$$

(if you do a row swap on $(A - \lambda I)$ before elimination).

Diagonalizing a matrix Once we have all n eigenvectors and eigenvalues of A we can put them into two matrices S and Λ, where S is the **eigenvector** matrix whose columns are $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ and Λ is a diagonal **Eigenvalue** matrix whose diagonal entries are the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. With these definitions we can always write

$$AS = S\Lambda \tag{31}$$

If S^{-1} exists, we can diagonalize A by

$$\Lambda = S^{-1}AS \tag{32}$$

or factor A into

$$A = S\Lambda S^{-1} \tag{33}$$

Rules for diagonalization

- A matrix can only be diagonalized if there are n independent eigenvectors (i.e. S is not singular)
- If all the eigenvalues are **different** then the eigenvectors are independent and A can be diagonalized.
- **Repeated eigenvalues** *might* prevent diagonalization (but not always)
- **Repeated eigenvectors always** prevent diagonalization

Symmetric Matrices $A = A^T$ In the special case that A is symmetric ($A = A^T$) then the following is always true

- All the eigenvalues of A are real.
- All the eigenvectors or A are orthogonal (and can be chosen orthonormal)
- **All** symmetric matrices can be diagonalized/factored as

$$A = Q\Lambda Q^T \tag{34}$$

  where Q is the eigenvector matrix (now chosen to be orthonormal).

Positive Definite Matrices For the special case of square symmetric matrices the special case where all the λ's are positive is called *positive definite*. If $\lambda >= 0$, the matrix is said to be *semi-definite*. A matrix will be positive definite if

1. All the pivots are positive
2. The number $\mathbf{x}^T A\mathbf{x} > 0$ for all vectors $\mathbf{x} \neq \mathbf{0}$

The two symmetric matrices $A^T A$ and AA^T are always at least semi-definite.

Applications of Diagonalization #1: Powers of matrices and Iterative maps Using the diagonalization theorem it is easy to show that the nth power of a diagonalizable matrix A is just

$$A^n = S\Lambda^n S^{-1} \tag{35}$$

this is useful for problems with iterative maps (e.g. the Fibonacci numbers) where a sequence of vectors is produced recursively by

$$\mathbf{x}_n = A\mathbf{x}_{n-1} \tag{36}$$

which can be unrolled to show that

$$\mathbf{x}_n = A^n\mathbf{x}_0 = S\Lambda^n S^{-1}\mathbf{x}_0 \tag{37}$$

Application #2: dynamical systems Eigenvectors and Eigenvalues are also very useful for solving dynamical systems such as

$$\frac{d\mathbf{u}}{dt} = A\mathbf{u} \tag{38}$$

Again, by decomposing **u** into its components in each direction of the eigenvectors (i.e.

$$\mathbf{u} = S\mathbf{y} \tag{39}$$

or $\mathbf{y} = S^{-1}\mathbf{u}$, we can transform Eq. (38) into

$$\frac{d\mathbf{y}}{dt} = \Lambda\mathbf{y} \tag{40}$$

which has the simple answer

$$\mathbf{y}(t) = e^{\Lambda t}\mathbf{y}_0 \tag{41}$$

where

$$e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \tag{42}$$

is a diagonal matrix, the matrix exponential of Λ. Transforming back into the standard basis using Eq. (39) gives

$$\mathbf{u}(t) = Se^{\Lambda t}S^{-1}\mathbf{u}_0 = e^{At}\mathbf{u}_0 \tag{43}$$

where

$$e^{At} = Se^{\Lambda t}S^{-1} = I + At + \frac{1}{2!}(At)^2 + \frac{1}{3!}(At)^3 \ldots \tag{44}$$

is the matrix exponential. An equivalent, but perhaps better way to write the solution (Eq. 43) is as

$$\mathbf{u}(t) = c_1e^{\lambda_1 t}\mathbf{x}_1 + c_2e^{\lambda_2 t}\mathbf{x}_2 + \ldots + c_ne^{\lambda_n t}\mathbf{x}_n \tag{45}$$

where the λ_i and $\mathbf{x}_i$ are each eigenvalue/eigenvector pair and c_i are the components of $\mathbf{c} = S^{-1}\mathbf{u}_0$ which is just the decomposition of the initial condition into the basis described by the eigenvectors.

Finally! the SVD The singular value decomposition (SVD) is another factorization of any general $m \times n$ matrix A into *two* orthogonal matrices and a diagonal matrix of *singular* values. The SVD of A is written

$$A = U\Sigma V^T \tag{46}$$

where U is orthogonal $m \times m$ ($U^TU = I$) and V is orthogonal $n \times n$ ($V^TV = I$) and

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \tag{47}$$

is a diagonal matrix[1] of singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$.

Some important properties of the SVD

- The columns of V are the eigenvectors of A^TA with eigenvalues σ_i^2. i.e. (for square A)

$$A^TA = V\Sigma^2V^T \tag{48}$$

  or $A^TA\mathbf{v}i = \sigma_i^2\mathbf{v}_i$
- The columns of U are the eigenvectors of AA^T also with eigenvalues σ_i^2
- If matrix A has rank r. The first r columns of V form an orthogonal basis for the **Row Space** of A, $C(A^T)$.
- the last $n - r$ columns form a basis for the **Null Space** of A, $N(A)$ (and the last $n - r$ σ_i's will be zero.)
- The first r columns of U are a basis for the **Column Space**, $C(A)$
- The last $m - r$ columns of U are a basis for the **Left Null Space** $N(A^T)$.

A recipe for finding the SVD

1. Form the symmetric matrix A^TA and find its eigenvalues λ_i.
2. Sort the eigenvalues from largest to smallest such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ and with this ordering factor A^TA into $A^TA = Q\Lambda Q^T$ where Λ is a diagonal matrix of *sorted* eigenvalues and Q is a square orthogonal matrix of the corresponding eigenvectors.
3. Then $V = Q$ and $\Sigma = \sqrt{\Lambda}$ (i.e. $\sigma_i = \sqrt{\lambda_i}$)

[1] Well close enough. If A is square, Σ is diagonal, otherwise it's as diagonal as an $m \times n$ matrix can get.

4. For all $\sigma_i > 0$ (i.e. $i = 1$ to r) find

$$\mathbf{u}_i = \frac{A\mathbf{v}_i}{\sigma_i}$$

(which should automatically be unit vectors)

5. If A has a left null space, there will be an additional $m - r$ vectors $\mathbf{u}_i$ for $i = r + 1, m$. For these just find an orthonormal basis for the left null space of A (i.e. find $N(A^T)$ then use Gram-Schmidt to orthogonalize).

Application of the SVD: Least squares and the pseudo inverse Given a singular value decomposition for an invertible matrix

$$A = U\Sigma V^T \tag{49}$$

the inverse of A is easily found to be

$$A^{-1} = V\Sigma^{-1}U^T \tag{50}$$

where

$$\Sigma^{-1} = \begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_n \end{bmatrix} \tag{51}$$

However, if A is singular then A^{-1} doesn't exist (because if A is singular, at least one of the $\sigma_i = 0$ and $1/\sigma_i \to \infty$. However it is easy to create the **Pseudo-inverse**

$$A^+ = V\Sigma^+U^T \tag{52}$$

where

$$\Sigma^+ = \begin{bmatrix} 1/\sigma_1 & & & & & \\ & \ddots & & & & \\ & & 1/\sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix} \tag{53}$$

i.e. the diagonal is simply $1/\sigma_i$ wherever that is defined and zero wherever it isn't. The best least squares solution for any problem $A\mathbf{x} = \mathbf{b}$ is then

$$\mathbf{x}^+ = A^+\mathbf{b} \tag{54}$$

In terms of the 4 subspaces, $AA^+\mathbf{b}$ is the projection of $\mathbf{b}$ onto the column space and $A^+A\mathbf{x}$ is the projection of $\mathbf{x}$ onto the row space.

6 Summary

Okay, here is a quick list of things to know

- Elimination, LU decomposition in all it's forms and solving square $A\mathbf{x} = \mathbf{b}$ for square A
- Solving $A\mathbf{x} = \mathbf{b}$ for general $m \times n$ A. (row reduced echelon form)
- Diagnosing the rank of A and the dimensions and bases for the four subspaces
- Projections onto lines and subspaces
- How to solve Least Squares Problems
- How to find orthogonal bases by Gram-Schmidt and factorizing $A = QR$
- Properties of Q matrices (and other special matrices such as elimination matrices, permutation, rotation, projection etc.)
- How to find the determinant of a matrix (and the properties of the determinant)
- How to find eigenvalues and eigenvectors of a matrix (and use tricks like the trace and determinant to check things)
- How to diagonalize a matrix and know when it's not possible
- Recognize the special properties of symmetric matrices
- Take powers of matrix and use the diagonalization theorem to solve dynamical systems.
- Find the SVD of a general matrix and use it to find the pseudo-inverse for singular least-squares problems
- Know various intriguing bits of Linear algebra trivia that help you do things faster.

7 P.S.

That's it for now...watch this space for anything new and/or corrections. if you have any questions come and see me in office hours or send me e-mail at mspieg@ldeo.columbia.edu to set up an appointment. Good luck and relax.