Math 214 (Winter '08)

Review for Midterm 2

March 14

Note: These notes are intended to summarize the essential material for Midterm 2. You still have to read the textbook and to review the recent HWs.

Chapter 4

- *Linear Space* (Def. 4.1.1). To have a linear space V you need to know how to add two elements of V and how to multiply by a scalar (constant). The addition and scalar multiplication satisfy the usual properties. Essential:

  (0) $0 \in V$;

  (1) closed under addition: $x, y \in V \Longrightarrow x + y \in V$;

  (2) closed under scalar multiplication: $x \in V \Longrightarrow c \cdot x \in V$.

  Concretely, to say that a certain set V is a linear (sub)space you have to explain that V is closed under addition and scalar multiplication (in our examples the meaning of addition and the scalar multiplication is obvious). Concrete examples: polynomials of certain degree $V = P_d$, square matrices $V = M_n$, functions with certain properties, etc. To test your understanding explain why the set V of upper diagonal 2×2 matrices form a linear space, but if I require in addition that the matrices in V are invertible, then V is no longer a linear space.

- *Linear Transformation* (Def. 4.2.1). To check that $T : V \to W$ is linear you have to check that T

  (1) behaves as expected w.r.t. addition: $T(x + y) = T(x) + T(y)$;

  (2) behaves as expected w.r.t scalar multiplication: $T(c \cdot x) = c \cdot x \in V$.

  In particular, one also have $T(0) = 0$. Typical examples: derivatives, multiplications by constant matrices, etc. Let $T : M_2 \to M_2$ and $A = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$. Explain why $T(X) = A \cdot X \cdot A$ is linear, but $T(X) = X \cdot A \cdot X$ is not.

- *Span, Linear independence, Basis, Dimension* (Def. 4.1.3). The key points to remember:

  – [basis] = [span] and [linear independent]

  – [dimension]= the number of elements in a basis

- in practice to check that something is a basis for V it is enough to check that you have the right numbers of elements and either they span V or that they are linearly independent.
- see (4.1.6) and the related examples for how to find a basis

Concrete examples: find a basis for the upper triangular matrices, for degree 2 polynomials passing through $(-2, 2)$, etc. Important examples find a basis for the image of a linear transformation, or for the kernel. Example: let $T : P_2 \to P_2$ be given by $T(f) = f'$, find a basis for $Ker(T)$ and $Im(T)$. Note the fundamental result:

$$\dim Ker(T) + \dim Im(T) = \dim V,$$

where V is the domain of T.

- *Isomorphism* (4.2.2 and 4.2.4). In practice the easiest way to see that $T : V \to W$ (a linear transformation) is an isomorphism is to check:
  - $\dim V = \dim W$
  - and $Ker(T) = 0$, i.e. solve the equation $T(x) = 0$ and deduce that the only solution is $x = 0$.

  Question: is it $T : P_2 \to P_2$ defined by $T(f) = xf'$ an isomorphism? What if $T(f) = xf' + f$?

- *Coordinates*. If we have given a basis $\mathcal{B} = \{v_1, \dots, v_n\}$ for V then we can transform any element x of V in vector $[x]_{\mathcal{B}}$ and work with as if we have $V = \mathbb{R}^n$ (the coordinate isomorphism). The role of this is to make the abstract notions of chapter 4, very concrete (as in the previous chapters). A few important points:
  - $$[x]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ \dots \\ a_n \end{bmatrix}$$

    simply means

    $$x = a_1 v_1 + \dots a_n v_n$$
  - with respect to the "standard basis" is quite obvious what is the coordinate vector. Say that $V = P_2$, $\mathcal{B} = \{x^2, x, 1\}$ and $f = 5x^2 + 3x - 7$. Then, clearly:

    $$[f]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 3 \\ -7 \end{bmatrix}$$

– if I try to do the same example as above, but with respect to the exotic basis $\mathcal{A} = \{x^2 + x + 1, x + 1, 1\}$ the situation becomes tricky. I need to find a, b, c such that

$$f = 5x^2 + 3x - 7 = a \cdot (x^2 + x + 1) + b \cdot (x + 1) + c \cdot 1$$

After expanding this becomes a linear system of equations. After solving, I get a, b, c and then

$$[f]_{\mathcal{A}} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

– Alternatively, since I can easily find $[f]_{\mathcal{B}}$, I can use the formula of change of basis:

$$[f]_{\mathcal{A}} = S_{\mathcal{B}\to\mathcal{A}} \cdot [f]_{\mathcal{B}}$$

to compute $[f]_{\mathcal{B}}$. The matrix $\mathcal{B} \to \mathcal{A}$ is obtained by taking the coordinates of the basis $\mathcal{B}$ with respect to the basis $\mathcal{A}$, i.e.

$$S_{\mathcal{B}\to\mathcal{A}} = [[v_1]_{\mathcal{A}}, \ldots, [v_n]_{\mathcal{A}}]$$

Concretely, in our example we express the basis $\mathcal{B} = \{x^2, x, 1\}$ in terms of the basis $\mathcal{A} = \{x^2 + x + 1, x + 1, 1\}$, i.e.

$$\begin{aligned} x^2 &= (x^2 + x + 1) - (x + 1) \\ x &= (x + 1) - 1 \\ 1 &= 1 \end{aligned}$$

Thus,

$$S_{\mathcal{B}\to\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

And, we get

$$[f]_{\mathcal{A}} = S_{\mathcal{B}\to\mathcal{A}} \cdot [f]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 3 \\ -7 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -10 \end{bmatrix}$$

(compare with the direct approach to compute $[x]_{\mathcal{A}}$).

– Note the important formula:

$$S_{\mathcal{B}\to\mathcal{A}} = (S_{\mathcal{A}\to\mathcal{B}})^{-1}$$

– Note also that $S_{\mathcal{A}\to\mathcal{B}}$ it is easy to find (compute it!).

- *The matrix of a linear transformation.* You have $T: V \to V$, and $\mathcal{B} = \{v_1, \ldots, v_n\}$ a basis. Then T is given (as in the case $V = \mathbb{R}^n$) by a $n \times n$ matrix B (s.t. $[T(x)]_{\mathcal{B}} = B \cdot [x]_{\mathcal{B}}$). The key formula is

  - $B = [[T(v_1)]_{\mathcal{B}}, \ldots, [T(v_n)]_{\mathcal{B}}]$, i.e. the columns of the matrix B are the coordinate vectors of the transforms (by T) of the basis vectors.

  Example: compute the matrix of the linear transformation $T: P_2 \to P_2$, where $T(f) = x \cdot f' + f$ with respect to the standard basis $\mathcal{B} = \{x^2, x, 1\}$. Now do the same with respect to the non-standard basis $\mathcal{A} = \{x^2 + x + 1, x + 1, 1\}$ (you should get another matrix A representing T). This can be done (and should be done for practice) in two ways:

  - the same type of computations as for $\mathcal{B}$
  - by using the base change formula:

    $$A = S \cdot B \cdot S^{-1},$$

    where $S = S_{\mathcal{B} \to \mathcal{A}}$ (N.B. S^{-1} can be computed as the inverse of S or by $S^{-1} = S_{\mathcal{A} \to \mathcal{B}}$).

Chapter 5

- *Orthogonal vectors, Orthonormal basis.* A few important points:

  - [orthonormal]=[ortho] (i.e. $u_i.u_j = 0$ if $i \neq j$) and [normal] (i.e. $||u_i|| = 1$);
  - if $\{u_1, \ldots, u_n\}$ are orthonormal, then they are linearly independent. Thus to check that $\{u_1, \ldots, u_n\}$ form an orthonormal basis for some linear subspace V it is enough to check that: (the vectors are orthonormal) and (the right number of vectors, i.e. $n = \dim V$).

- With respect to orthonormal bases it is easy to compute coordinates and projections. Let V a linear subspace (in some $\mathbb{R}^N$) and $\mathcal{B} = \{u_1, \ldots, u_n\}$ an orthonormal basis. The following two formulas are essential:

  - if $x \in V$, then

    $$[x]_{\mathcal{B}} = \begin{bmatrix} u_1.x \\ \ldots \\ u_n.x \end{bmatrix}$$

    (the coordinates are computed by taking the dot product).
  - For any x (not necessary in V), the orthogonal projection on V is computed by

    $$x^{||} = proj_v(x) = (x.u_1)u_1 + \ldots (x.u_n)u_n.$$

Note also that

$$x = x^{||} + x^{\perp}$$

In the Gram-Schmidt algorithm, you need $x^{\perp}$, which is computed by $x^{\perp} = x - x^{||}$ and the previous formula.

- *Gram-Schmidt* (see 5.2). You are given an arbitrary basis $\{v_1, \ldots, v_n\}$ for V and want to find an orthonormal basis $\{u_1, \ldots, u_n\}$. The idea is quite simple: it easy to get norm 1 (just normalize the vectors), the hard part is to obtain "ortho"; this is done by computing the orthogonal projection. Specifically:

(Step 1) Normalize v_1, i.e. $u_1 = \frac{v_1}{||v_1||}$

(Step 2) Compute $v_2^{\perp}$ as explained above, i.e.

$$v_2^{\perp} = v_2 - (v_2.u_1)u_1$$

(Step 3) Normalize $v_2^{\perp}$, i.e. $u_2 = \frac{v_2^{\perp}}{||v_2^{\perp}||}$

(Step 4) Compute $v_3^{\perp}$ (w.r.t. u_1, u_2)

$$v_3^{\perp} = v_3 - (v_3.u_1)u_1 - (v_3.u_2)u_2$$

(Step 5) Normalize $v_3^{\perp}$

(etc.)

- To compute an orthogonal basis for a linear space V (e.g. $Ker(A)$, $Im(A)$), you first need to find a basis $\{v_1, \ldots, v_n\}$ (as we did in the previous chapters), and then apply Gram-Schmidt to get an orthonormal basis. Recall that to find a basis for $Im(A)$ you simply have to remove the redundant vectors among the columns of A. To find a basis for $Ker(A)$ you have to find the relations between the columns vectors of A.

- Suppose that you are given a linear subspace V, then the orthogonal complement $V\perp$ is defined as the set of vectors orthogonal on V. To find $V^{\perp}$ you have to solve the linear system

$$\begin{aligned} x.v_1 &= 0 \\ &\cdots \\ x.v_n &= 0 \end{aligned}$$

where $v_1, \ldots, v_n$ is a basis for V. The orthogonal complement has complementary dimension to V, i.e.

$$\dim V + \dim V^{\perp} = N$$

where N is the dimension of the ambient space (V is a subspace in some $\mathbb{R}^{\mathbb{N}}$). Example: Let $V = Span\left(\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}4\\5\\6\end{bmatrix}\right)$. Then $V^{\perp}$ is 1-dimensional (since $\dim V = 2$ and $\dim V^{\perp} + \dim V = 3$) and $V^{\perp}$ contains all the vectors $\begin{bmatrix}a\\b\\c\end{bmatrix}$ orthogonal to $\begin{bmatrix}1\\2\\3\end{bmatrix}$ and $\begin{bmatrix}4\\5\\6\end{bmatrix}$. This means

$$\begin{aligned} a + 2b + 3c &= 0 \\ 4a + 5b + 6c &= 0 \end{aligned}$$

You have to solve for a, b, c (i.e. express a, b in terms of the free variable c). Since $V^{\perp}$ is 1-dimensional, a basis is given by a non-trivial solution of the previous linear system (just set the free variable c to some random non-zero value).

- In fancy words, in the previous example $V = Im(A)$ and $V^{\perp} = Ker(A^T)$ (see 5.4.1), where

$$A = \begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix} \quad \text{and} \quad A^T = \begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}.$$

This means that V is spanned by the column vectors of A, and $V^{\perp}$ is obtained by solving the linear system $A^T \cdot x$ (i.e. the linear system considered a few lines above).

- *Orthogonal Transformations/Matrices.* A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if it preserves the norm: $||T(x)|| = ||x||$. An orthogonal transformation automatically preserves also the angles, i.e. $x \perp y \implies T(x) \perp T(y)$. This gives an alternative characterization of orthogonal transformations:

  – T is orthogonal if and only if $\{T(e_1), \ldots T(e_n)\}$ forms an orthonormal basis.

  The same thing in terms of the matrix A representing T is:

  – A is orthogonal if and only if the columns of A form an orthogonal basis

  The important fact about orthogonal matrices is that it is easy to invert:

  – $A^{-1} = A^T$, where A^T is the transpose of A (the columns of A^T are the rows of A).

- Least square In many situations we are not able to solve precisely the linear system:

$$Ax = b$$

In fact, the linear system can be solved only if $b \in Im(A)$. To find an approximate solution, we project b onto $V = Im(A)$. Namely, $Ax = proj_V b$ has always a solution x^* and this solution is optimal, in the sense that it minimizes the error:

$$\text{Error} = ||b - Ax^*|| = ||b - proj_V b||$$

- Concretely, to get the best approximate solution to a linear system $Ax = b$, we solve the linear system

$$A^T A x^* = A^T b$$

This system has always a solution x^*, which is the optimal approximate solution. Note also

$$proj_V b = Ax^* = A(A^T A)^{-1} A^T b$$

(if $Ker(A) = 0$). If A is orthogonal, the formula simplifies to

$$proj_V b = Ax^* = AA^T b.$$