Study Guide, Part II

Linear transformations

Definition: $T : V \to W$ is a *linear transformation* if T is compatible
with addition, scalar multiplication

Often we have $W = V$, so $T : V \to V$
- In this case, T is sometimes called a *linear operator*

Examples
- rotation in $\mathbf{R}^2$
- reflection in $\mathbf{R}^2$
- isomorphism $V \cong W$
- a "matrix transformation" T_A (see below)
- differentiation, for example on $P(2) \to P(1)$
  - also OK to say $P(2) \to P(2)$, or $P(n) \to P(n)$
    - also for solution spaces of linear DE's, such as $y'' = -y$.

Associated subspaces
- image (range)
  - is set of values of T
  - same: $\{T(\mathbf{v}) : \mathbf{v} \in V\}$
  - same: set of all $\mathbf{w} \in W$ with $\mathbf{w} = T(\mathbf{v})$ for some $\mathbf{v} \in V$
- kernel
  - is set of vectors $\mathbf{v}$ with $T(\mathbf{v}) = \mathbf{0}$

T_A as linear transformation
- definition
  - for a matrix A, say 2×2, define the "matrix transformation" $T_A : \mathbf{R}^2 \to \mathbf{R}^2$ by $T_A(\mathbf{x}) = A\mathbf{x}$
- how to visualize–e.g., the house pictures
- image of T_A = span of columns of A = column space of A
- kernel of T_A = null space of A
- (Also, the row space of A can be expressed using T_A; see below.)

Matrix of a linear transformation $T : V \to W$ relative to a basis B of V and a basis C of W. Say B is $\mathbf{v}_1, \ldots, \mathbf{v}_n$

Notation in text: $[T : B, C]$

Column $\#i$ is the coordinate vector of $T(\mathbf{v}_i)$ relative to C.

Be able to find this matrix in situations where V and W are spaces of functions and T is differentiation, for example.

Square matrices

Special matrices

I

Zero matrix

scalar matrices

diagonal matrices

how to add, multiply them, take powers

T_A scales axes, separately

Determinant of A:

Absolute value of det A tells

the factor by which T_A changes all areas, if A is 2×2, so $T_A : \mathbf{R}^2 \to \mathbf{R}^2$

or volumes, if A is 3×3, so $T_A : \mathbf{R}^3 \to \mathbf{R}^3$.

Sign of det A tells whether orientation is preserved

$\det A = 0 \Leftrightarrow A$ is singular.

There is a version for $n \times n$ matrices.

Determinants make sense only for square matrices!

Nonsingularity

equivalent conditions (all true about A, or all false)

linear equations with A as coeff are solvable, uniquely

$\det \neq 0$

A is invertible, so have $AA^{-1} = I$, $A^{-1}A = I$.

A has rank n (the largest possible)

T_A is onto

T_A is one-to-one

T_A is an isomorphism

Three uses of square matrices:

1. $T_A : \mathbf{R}^n \to \mathbf{R}^n$ ($\mathbf{R}^n$ itself)

2. Change-of-basis matrix (transition matrix)

   Expresses new basis relative to old

3. Matrix of a linear transformation from a space to itself:

   $T : V \to V$, with respect to one basis

   Notation in text: $[T : B]$

   What if change basis in V?

   Suppose matrix for old basis is A, transition matrix P.

   Get new matrix $P^{-1}AP$.

   What is the more general statement if have $T : V \to W$ with bases B, C, so that A is not necessarily square?

The theory of eigenvalues of square matrices deserves a separate section:

Eigenvalues (and eigenvectors) of square matrices

characteristic polynomial

how to find with trace, det if $n = 2$.

eigenvalues

multiplicity

sum = trace A, product = det A

0 is an eigenvalue $\Leftrightarrow$ A is singular

easy cases: diagonal matrices, triangular matrices

In many problems, knowing the eigenvalues already gives valuable information, for example the rate of growth or decay, or the frequency of oscillation.

eigenvectors

must be nonzero, by definition!

any nonzero scalar times an eigenvector is an eigenvector

eigenspace E_λ

consists of all eigenvectors for λ and also $\mathbf{0}$

$E_\lambda = \{\mathbf{v} : T(\mathbf{v}) = \lambda\mathbf{v}\}$

E_λ is a subspace

When λ is not an eigenvalue, E_λ is defined but is $\{\mathbf{0}\}$.

Application of eigenvalues to stochastic matrices similarity

- preserves det
- preserves characteristic polynomial
- preserves eigenvalues
- compatible with matrix powers
- in fact, compatible with matrix polynomials, e.g., $A \sim B \Rightarrow A^2 + A + I \sim B^2 + B + I$.
- Cayley-Hamilton, use in inventing examples
- Get a similarity when take $T_A : \mathbf{R}^n \to \mathbf{R}^n$ and change the basis

diagonalization

- what's good about diagonal matrices
- "diagonalizing" A means finding P with $P^{-1}AP = D$
  - diagonal entries of D are the eigenvalues
  - columns of P are the corresponding eigenvectors
    - We are really choosing new basis so matrix of lin. tr. is diagonal
- A is diagonalizable $\Leftrightarrow$ there are n LI eigenvectors
- Defective matrix is opposite of diagonalizable: there are $< n$ LI eigenvectors, even if complex numbers are allowed
- Cases guaranteed to be diagonalizable
  - distinct eigenvalues
  - real symmetric
- Examples that are defective (in other words, not diagonalizable):
  - a nilpotent nonzero matrix N
  - a matrix that has just one eigenvalue but is *not* a scalar matrix
- Uses of diagonalization:
  - Often turns a difficult problem into an easy one.
  - Use in solving linear systems of DE's, $\mathbf{x}' = A\mathbf{x}$
    - Substitute $\mathbf{r}$ so that $\mathbf{x} = P\mathbf{r}$
    - Get diagonal system, solve, put back in terms of $\mathbf{x}$
    - Also, if initial vector is eigenvector of A, get a "pure", exponential solution
  - Lab: Used to solve $\mathbf{x}'' = A\mathbf{x}$
    - Not responsible for writing out solution of second-order systems, but good practice to try solving anyway using a
      substitution.

Inner products and orthogonality

Terminology

"inner product" = "dot product" = "scalar product"

"orthogonal" = "perpendicular"

Beware: the term "orthogonal matrix" means a matrix with *orthonormal* columns

inner product $\langle \mathbf{v}, \mathbf{w} \rangle$ in $\mathbf{R}^n$

can express

algebraically as $v_1 w_1 + \cdots + v_n w_n$

(where v_i, w_i are entries, not separate vectors)

geometrically as $\|\mathbf{v}\|\|\mathbf{w}\| \cos\theta$

as matrix product $\mathbf{v}^T \mathbf{w}$

$\cos\theta = \cdots$

$\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v} \rangle$

$\mathbf{v} \perp \mathbf{w} \Leftrightarrow \langle \mathbf{v}, \mathbf{w} \rangle = 0$

projection of $\mathbf{v}$ on the line in the direction of **unit vector u**

scalar projection is $\langle \mathbf{v}, \mathbf{u} \rangle$

vector projection is $\langle \mathbf{v}, \mathbf{u} \rangle \mathbf{u}$

can express $\mathbf{v}$ as sum of two components

one in direction of $\mathbf{u}$ (the vector projection)

one perpendicular to $\mathbf{u}$ ($\mathbf{v}$ minus the vector projection)

So it's $\mathbf{v} - \langle \mathbf{v}, \mathbf{u} \rangle \mathbf{u}$.

Can regard perpendicular component as obtained by "straightening up" $\mathbf{v}$ relative to $\mathbf{u}$.

If have a nonzero vector $\mathbf{w}$ for the direction instead of a *unit* vector $\mathbf{u}$,

put $\mathbf{w}/\|\mathbf{w}\|$ in place of u in formulas,

get perpendicular component $\mathbf{v} - \dfrac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w}$

orthonormal list of vectors

$\mathbf{v}_1, \ldots, \mathbf{v}_n$ are *orthonormal* if

(1) $\|\mathbf{v}_i\| = 1$ for each i and

(2) $\mathbf{v}_i \perp \mathbf{v}_j$ for $i \neq j$.

Here "ortho" = "orthogonal" and "normal" = "length 1".

Same: $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 1$ if $i = j$, or $= 0$ if $i \neq j$.

To *orthogonalize* a list of nonzero vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ means to find mutually orthogonal vectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$ with the same span as $\mathbf{v}_1, \ldots, \mathbf{v}_n$

Gram-Schmidt orthogonalization method

$\mathbf{w}_1 = \mathbf{v}_1$;

$\mathbf{w}_2$ is obtained by straightening up $\mathbf{v}_2$ relative to $\mathbf{w}_1$;

$\mathbf{w}_3$ is obtained by straightening up $\mathbf{v}_3$ relative to $\mathbf{w}_1$ and then $\mathbf{w}_2$,

etc.

Equivalently, $\mathbf{w}_i$ is $\mathbf{v}_i$ minus the component of

$\mathbf{v}_i$ in the span of $\mathbf{w}_1, \ldots, \mathbf{w}_{i-1}$.

Proof that span is the same: Put $\mathbf{v}_1, \ldots, \mathbf{v}_n$ as rows of a matrix, notice that $\mathbf{w}_1, \ldots, \mathbf{w}_n$ come from doing elementary row operations, so row space is the same.

ortho*normalization*

means finding orthonormal vectors

$\mathbf{u}_1, \ldots, \mathbf{u}_n$ with the same span as

$\mathbf{v}_1, \ldots, \mathbf{v}_n$:

Apply Gram-Schmidt to get $\mathbf{w}_1, \ldots, \mathbf{w}_n$, but

scale the vectors $\mathbf{w}_i$ to get unit vectors $\mathbf{u}_i$, either

one-by-one, as soon as each $\mathbf{w}_i$ is computed, or

together, after finding all of $\mathbf{w}_1, \ldots, \mathbf{w}_n$.

orthogonal subspaces [if we get to this in lecture]

If W is a subspace of $\mathbf{R}^n$, $W^\perp$ is the set of all vectors orthogonal to all vectors in W.

$W^\perp$ is also a subspace, the *orthogonal complement* of W.

$(W^\perp)^\perp = W$.

Example:

For a matrix A, vectors in row space are $\perp$ vectors in null space.

Can regard row space as $N(A)^\perp$, where $N(A)$ is null space.

row space of $A = (\ker T_A)^\perp$,

$\ker T_A = \text{row space}^\perp$.

orthogonal matrices

- equivalent characterizations
  - the definition (orthonormal columns)
  - $P^T P = I$
  - P is invertible and $P^{-1} = P^T$
  - P^t is orthogonal (i.e., rows of P are orthonormal)
  - T_P preserves inner products
  - T_P is rigid
- The inverse (=transpose) of an orthogonal matrix is orthogonal.
- The product of two orthogonal matrices is orthogonal.
- $\det P = \pm 1$
  - Definition: A *rotation* matrix is an orthogonal matrix with determinant 1 (rather than -1)

Symmetric matrices are closely tied to inner products:

- A is symmetric $\Leftrightarrow \langle A\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{v}, A\mathbf{w}\rangle$

Spectral theorem: If A is a real symmetric matrix, then

- A has real eigenvalues,
  - Eigenvectors for two different eigenvalues are perpendicular.
- A can be diagonalized, using a rotation matrix
  - (There are fancier versions of the spectral theorem in advanced courses.)

Application to quadratic forms

- Given quadratic form such as $2x^2 + 6xy + 7y^2$,
  - write as $\mathbf{x}^t A\mathbf{x}$
  - Diagonalize as $R^{-1}AR = D$
  - Substitute $\mathbf{x} = R\mathbf{r}$
  - Get quadratic form $\lambda_1 r^2 + \lambda_2 s^2$ with no "cross term" rs.
  - Now clear what the graph of $2x^2 + 6xy + 7y^2 = 1$ is.