Review and True/False questions, Bretscher 2.2-3.1

Review

2.2 Linear transformations in geometry.

Before proceeding, remember that we use notation $e_1, \dots, e_n$ for the standard basis, i.e.

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Also, recall, that for any matrix $A_{m\times n}$ one has

$$A = \begin{bmatrix} | & & | \\ Ae_1 & \dots & Ae_n \\ | & & | \end{bmatrix}$$

We covered what different 2x2 matrices mean geometrically, accompanying each case with illustrations.

1. *Scaling.* Matrix $A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$ corresponds to scaling the plane by a factor k. I.e. $A\vec{x} = k\vec{x}$ for any vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

2. *Projection.* Suppose $a^2 + b^2 = 1$. Then matrix $A = \begin{bmatrix} a^2 & ab \\ ab & b^2 \end{bmatrix}$ corresponds a projection on the line $L = \{t \begin{bmatrix} a \\ b \end{bmatrix} | t \in \mathbb{R}\}$, which is the line "along" the vector $\begin{bmatrix} a \\ b \end{bmatrix}$.

3. *Reflection.* Suppose $a^2+b^2 = 1$. Then matrix $A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$ corresponds to a reflection about the line $L = \{t \begin{bmatrix} a \\ b \end{bmatrix} | t \in \mathbb{R}\}$, which is the line "along" the vector $\begin{bmatrix} a \\ b \end{bmatrix}$.

4. *Rotation.* Suppose $a^2 + b^2 = 1$ (one can write them as $a = \cos\theta$ and $b = \sin\theta$). Then matrix $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ corresponds to a counter-clockwise rotation

of the plane around the origin by the angle θ, which we denote by R_θ. Note, that $R_\theta(e_1) = Ae_1 = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$.

5. *Rotation+scaling.* For any a, b matrix $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = r \cdot \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ (where $r^2 = a^2 + b^2$ and $r \geqslant 0$) corresponds to a counter-clockwise rotation of the plane around the origin by the angle θ, composed with scaling by a factor r. Note, that $Ae_1 = \begin{bmatrix} a \\ b \end{bmatrix} = r \cdot \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$.

6. *Shears.* Matrices $S_h = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}, S_v = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$ correspond to horizontal and vertical sheers respectively. In the case of horizontal sheer, the image of y-axis is a line of a slope $1/k$, passing through origin. Points on x-axis are preserved by A.

2.3 Matrix multiplication

Definition 1. If $A_{k\times m} = \begin{bmatrix} | & & | \\ \vec{v}_1 & \dots & \vec{v}_m \\ | & & | \end{bmatrix}$ and $B_{l\times k} = \begin{bmatrix} - & \vec{w}_1 & - \\ & \dots & \\ - & \vec{w}_l & - \end{bmatrix}$ then

$$B_{l\times k} \cdot A_{k\times m} = \begin{bmatrix} - & \vec{w}_1 & - \\ & \dots & \\ - & \vec{w}_l & - \end{bmatrix} \cdot \begin{bmatrix} | & & | \\ \vec{v}_1 & \dots & \vec{v}_m \\ | & & | \end{bmatrix} = \begin{bmatrix} \vec{w}_1 \cdot \vec{v}_1 & & \vec{w}_1 \cdot \vec{v}_m \\ & \ddots & \\ \vec{w}_l \cdot \vec{v}_1 & & \vec{w}_l \cdot \vec{v}_m \end{bmatrix}$$

Important. Matrix multiplication corresponds to a composition of linear transformations. (illustration)

Properties of matrix multiplication:

1. $(AB)C = A(BC)$
2. $A(B + C) = AB + AC$
3. $(kA)B = k(AB)$
4. AB does not always equal to BA!

Practically, this means that one should treat multiplication of matrices as multiplication of numbers, except often $AB \neq BA$. (another key difference is that one cannot always divide by a matrix, see the next section).

Example 1. We illustrated $AB \neq BA$ on the following two matrices:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

First, we just multiplied them both ways and saw that the results are not the same. Second, we understood geometrically why $AB \neq BA$, by looking at the image of the frame (e_1, e_2) under the linear transformations AB and BA.

2.4 Inverse

Recall that identity matrix $Id_{n\times n} = I_n$ is a matrix which has 1's on its diagonal, and 0's everywhere else. This matrix can be characterized as the only linear transformation $A : \mathbb{R} \to \mathbb{R}$, s.t. $A\vec{x} = \vec{x}$ for any $\vec{x} \in \mathbb{R}$.

Definition 2. Linear transformation $T_1 : \mathbb{R}^n \to \mathbb{R}^n$ is invertible if it is invertible as a function, i.e. there is such a linear transformation $T_2 : \mathbb{R}^n \to \mathbb{R}^n$, that $T_2(T_1(\vec{x})) = \vec{x}$ for any $\vec{x} \in \mathbb{R}^n$.

Notice, that from this definition it is clear, that the matrix which represents composition $T_2 \circ T_1$ is I_n.

There is a corresponding notion of an inverse for matrices.

Definition 3. Matrix $B_{n\times n}$ is said to be inverse of the matrix $A_{n\times n}$, if $AB = I_n$.

Definition 4. Matrix $B_{n\times n}$ is said to be inverse of the matrix $A_{n\times n}$, if $BA = I_n$.

These definitions are equivalent. Thus, it is may be better to remember the following definition:

Definition 5. Matrix $B_{n\times n}$ is said to be inverse of the matrix $A_{n\times n}$, if $BA = AB = I_n$. In this case we denote $B = A^{-1}$.

If one has an equation $A\vec{x} = \vec{b}$, and A is invertible (i.e. has an inverse), then one can find a solution by multiplying both sides by A^{-1}:

$$A^{-1}A\vec{x} = A^{-1}\vec{b} \iff \vec{x} = A^{-1}\vec{b}$$

($\iff$ means implication in both directions, and reads as "equivalently")

Thus it is important to understand: 1) Does the matrix has an inverse? 2) How one can find an inverse?

Criterion. *A is invertible $\iff$ A is a square matrix of size $n \times n$ and it has the full rank $rk(A) = n$. This is equivalent to saying that $rref(A) = I_n$, because $rk(A)$ is, by definition, the number of leading ones in rref.*

Theorem (how to find inverse)**.** *If $A_{n\times n}$ is invertible, then one can use Gauss-Jordan elimination method to find A^{-1}. Namely, one has to find rref of $n \times 2n$ matrix $[A|I_n]$:*

$$rref\Big(\Big[\ A \mid I_n\ \Big]\Big) = \Big[\ I_n \mid A^{-1}\ \Big]$$

Note also, that if A and B are invertible matrices, we have $(AB)^{-1} = B^{-1}A^{-1}$. Note that the order of A and B changes.

3.1 Image and Kernel

This is one of the most important sections in the book, it is important to understand well these two concepts.

Suppose we have a linear transformation, given by a matrix $A_{n\times m}$:

$$A : \mathbb{R}^m \to \mathbb{R}^n$$

Definition 6. Image of A is a subset of a target $Im(A) \subset \mathbb{R}^n$, consisting of such elements $\vec{y} \in \mathbb{R}^n$, that $A\vec{x} = \vec{y}$ for some $\vec{x} \in \mathbb{R}^m$ in the domain.

Definition 7. Kernel of A is a subset of a domain $Ker(A) \subset \mathbb{R}^m$, consisting of such elements $\vec{x} \in \mathbb{R}^m$, that $A\vec{x} = 0$.

Important. Finding Kernel of A is the same as solving an equation $A\vec{x} = 0$.

Definition 8. Suppose we have vectors $v_1, v_2, \ldots, v_k \in \mathbb{R}^m$. Then $span(v_1, v_2, \ldots, v_k) \subset \mathbb{R}^m$ is the set of vectors consisting of all possible linear combinations $c_1v_1 + \cdots + c_kv_k$.

Statement. *Suppose* $A = \begin{bmatrix} | & & | \\ v_1 & \ldots & v_k \\ | & & | \end{bmatrix}$ *(by the way, remember that this means* $v_i = Ae_i$*, see the beginning of these notes). Then one has*

$$Im(A) = span(v_1, v_2, \ldots, v_k)$$

How to write Ker(A) as a span? Finding kernel is the same as solving $A\vec{x} = 0$. So just use the Gauss-Jordan elimination method to solve it, and in the end you will get a parameterization of the space of solutions by free variables. From this one can write Ker(A) as a span. (and therefore, as an image)

How to write Im(A) as a kernel? This one is trickier, we won't cover it, but it is extremely useful to know how to do that. Key idea: first find (write as a span or image) the orthogonal compliment of Im(A) by solving $A^T\vec{x} = 0$. Then one can write $Ker(A^T) = Im(B)$. Then $Ker(B)$ is the one you were looking for.

Theorem. *Suppose A is a square* $n \times n$ *matrix. Then the following are equivalent:*

1. *A is invertible.*
2. $Ker(A) = 0$*, i.e. the only vector sent to 0 by A is 0.*
3. $Im(A) = \mathbb{R}^n$*, i.e. the image is the whole* $\mathbb{R}^n$*.*
4. $rank(A) = n$ *(remember that this is equivalent to* $rref(A) = I_n$*).*
5. $\det(A) \neq 0$ *(this one we will cover later)*

True/False questions

First we covered a couple of T/F questions from previous review session (chapters 1.1-2.1 of the book) material.

Then we discussed what is "proof by contradiction", see Example 3 from "T:F_intro.pdf" from last time, or *wikipage*.

Warning: "proof by contradiction" is not the same as "proof by contrapositive", which is discussed in Appendix B of Bretscher. "Proof by contradiction" is a general strategy for proofs of any statements. "Proof by contrapositive" can only be applied to statements like "A => B". The strategy is to prove instead "not B => not A".

T/F questions for this time:

Problem 1. *There exists an invertible matrix $A \neq I$ such that $A^2 = A^3$.*

Solution. False. Let us prove it by contradiction. Suppose there is such an invertible matrix A. Then, by multiplying both sides of $A^2 = A^3$ by A^{-2} (here we use invertibility), one gets $I = A$. This contradicts the assumption $A \neq I$. Therefore such an invertible matrix does not exists. $\square$

Problem 2. *Any reflection is invertible.*

Solution. True. Reflection is an inverse of itself. $\square$

Problem 3. *A is $n \times n$. Then $A^2 = 0$ if and only if $Im(A) \subseteq Ker(A)$*

Solution. True. One has to write down the definition of Im(A) and Ker(A). $\square$

Remark. Notice, that if one wants to prove statement like "A is if and only if B", then he really needs to prove two statements. One is "A => B", and another is "B => A".

Problem 4. *There is a 2×2 matrix A such that Im(A)=Ker(A).*

Solution. True. Matrix $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is an example. $\square$

Problem 5. *If A is $n \times n$ matrix with $A^2 = 0$, then $I + A$ is invertible.*

Solution. True. $I - A$ is the inverse. Also, one can show directly that there is no such $\vec{x} \neq 0$, that $(I + A)\vec{x} = 0$ (cause otherwise $A(\vec{x}) = -\vec{x}$ and $A^2\vec{x} = \vec{x}$). This means that Ker(I+A)=0, which means I+A is invertible by a theorem above. $\square$

Problem 6. *If A,B are invertible and of the same size, then A+B is invertible. Harder version – assume also that A,B have no negative entries.*

Solution. False. Take $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. $\square$

Review and True/False questions, Bretscher 3.2-3.3

Review

3.2 Subspaces, bases, linear dependence

Definition 1. Vectors $\vec{v}_1, \ldots, \vec{v}_n$ are called linearly independent if there are no "relations" between them, i.e. there is no non-trivial linear combination $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$ which is equal to zero. In other words, if $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = 0$, then one has to have $c_1 = \cdots = c_n = 0$.

Example 1. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ are lin. independent.

Example 2. $\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ are lin. dependent.

Definition 2. Subspace of $\mathbb{R}^n$ is a subset $W \subset \mathbb{R}^n$ s.t.

1. $\vec{0} \in W$
2. $\vec{x} \in W \implies c\vec{x} \in W$ for any scalar c (i.e. real number). This is being closed under multiplication.
3. $\vec{x}, \vec{y} \in W \implies \vec{x} + \vec{y} \in W$. This is being closed under addition.

The point is that subspaces behave just like linear spaces $\mathbb{R}^k$, they just sit inside a bigger R^n. I

Example 3. We listed all the subspaces in $\mathbb{R}^1, \mathbb{R}^2$.

Example 4. Any $span(\vec{v}_1, \ldots, \vec{v}_n)$, any $Im(A)$, and any $Ker(A)$ are subspaces.

Fact. Set of solutions to $A\vec{x} = \vec{b}$ is a subspace only if $\vec{b} = 0$ (illustrated on the blackboard).

Definition 3. Set of vectors $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is called a basis of a subspace $W \subset \mathbb{R}^n$ if

1. $span(\vec{v}_1, \ldots, \vec{v}_n) = W$
2. $\vec{v}_1, \ldots, \vec{v}_n$ are lin. independent

Every subspace has a basis. It is not unique (illustrated on the blackboard). But the number of vectors in the basis is unique, see the next section.

Important (finding bases of Im(A) and Ker(A)). We illustrated in details two methods below on a matrix

$$A = \begin{bmatrix} 2 & 4 & 6 & 18 \\ 1 & 2 & 4 & 3 \end{bmatrix}$$

1. How to find a basis of $span(\vec{v}_1, \ldots, \vec{v}_k)$? In other words, how to eliminate all redundant vectors? In other words, how to find basis of $Im(A)$, where $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \ldots & A\vec{v}_k \\ | & & | \end{bmatrix}$?

   The answer is to just run Gauss-Jordan elimination method to find $rref(A)$, and then pick into basis of $Im(A)$ those columns of A, which correspond to leading ones in $rref(A)$. This method works, because elementary row operations do not change relations between column vectors. The columns with leading ones are linearly independent by Theorem 3.2.5 from the Bretscher.

2. How to find a basis of $Ker(A)$? One again has to do Gauss-Jordan elimination. After that one has a parameterization of the space of solutions to $A\vec{x} = 0$ by free variables. Treating free variables as coefficients in the linear combination, one gets a presentation of $Ker(A)$ as a span of vectors. Those vectors will be the basis of $Ker(A)$.

Remark. These two methods of finding basis actually prove rank-nullity theorem, because $dim(Ker(A)) = nullity(A)$ is a number of free variables in $rref(A)$, and $dim(Im(A)) = rk(A)$ is the number of leading one in $rref(A)$.

3.3 Dimension of a subspace

Definition 4. Dimension of a subspace $W \subset \mathbb{R}^n$ is the least number of vectors in the basis of W.

Fact (definition of dimension is correct)**.** Any subspace $W \subset \mathbb{R}^n$ has a basis. It is not unique, but the number of vectors in it is always the same.

Example 5. Find out what is the dimension of a line, plane, of our space, of the whole space-time universe ($\mathbb{R}^4$). What about general case of $\mathbb{R}^n$? (rmk: the dimension of $\mathbb{R}^0$ is 0)

Example 6. Find out what is the dimension of every subspace of $\mathbb{R}^2$, $\mathbb{R}^3$ (simultaneously recalling what are those subspaces). Notice, that it makes sence that subspace always has less dimension then the ambient space.

Now we are going to state theorems, which are important, and which we are going to use while solving T/F questions.

Theorem (how to tell if vectors are linearly independent)**.** *Suppose we are given m vectors $\vec{v}_1, \ldots, \vec{v}_m \in \mathbb{R}^n$. Then the following are equivalent (i.e. if one of the statements is true, the others should also true):*

1. $\vec{v}_1, \ldots, \vec{v}_m$ are linearly independent

2. there is no redundant vector in $\vec{v}_1, \ldots, \vec{v}_m$, i.e. $\vec{v}_1, \ldots, \vec{v}_m$ give a basis of $span(\vec{v}_1, \ldots, \vec{v}_m)$.

3. *none of $\vec{v}_i$ is a linear combination of the rest of the vectors*

4. *Ker* $\begin{bmatrix} | & & | \\ \vec{v}_1 & \dots & \vec{v}_m \\ | & & | \end{bmatrix} = \{\vec{0}\}$

5. *rank* $\begin{bmatrix} | & & | \\ \vec{v}_1 & \dots & \vec{v}_m \\ | & & | \end{bmatrix} = m$

The next theorem is one of the main ones.

Theorem (rank-nullity theorem)**.**
(Number of columns of A)=dim(Ker(A)) + dim(Im(A)).

It is called rank-nullity theorem because dim(Ker(A)) is called nullity(A), and dim(Im(A))=rk(A). Let's illustrate rank-nullity theorem on goto matrices for reflection, projection, rotation.

Example 7. Reflection $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (general form is $A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$ having $a^2 + b^2 = 1$). The rank is 2, while the nullity is 0.

Example 8. Projections: $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. The ranks are 1 and 2, the nullities are 1 and 1 respectively.

Example 9. Rotation $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (general form is $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$). The rank is 2, the nullity is 0.

This is a very useful fact to remember:

Fact (rank inequality for the product)**.**
For any two matrices (which one can multiply) the following inequalities are true:

$$rk(A \cdot B) \leqslant rk(A)$$

$$rk(A \cdot B) \leqslant rk(B)$$

Finally, let us repeat the invertibility criterion, now adding two more equivalent statements (last ones), and adding to the statement the fact that rk(A)=dim(Im(A)).

Theorem (different disguises of invertibility)**.** *Suppose A is a square $n \times n$ matrix. Then the following are equivalent:*

1. *A is invertible.*

2. *$Ker(A) = 0$, i.e. the only vector sent to 0 by A is 0.*

3. *$Im(A) = \mathbb{R}^n$, i.e. the image is the whole $\mathbb{R}^n$.*

4. $rk(A) = dim(Im(A)) = n$ *(remember that this is equivalent to* $rref(A) = I_n$*).*

5. $\det(A) \neq 0$ *(this one we will cover later)*

6. All rows of A are linearly independent.

7. All columns of A are linearly independent.

True/False questions

Problem 1. $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}$. If $\vec{v}_1, \vec{v}_2$ are linearly independent, and $\vec{v}_2, \vec{v}_3$, then $\vec{v}_1, \vec{v}_3$ are linearly independent.

Solution. False. Counterexample would be $\vec{e}_1, \vec{e}_2, \vec{e}_1$ on a plane. □

Problem 2. Set of vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ in $\mathbb{R}^2$ satisfying $x^2 = y^2$ is a subspace.

Solution. False. It is a union of lines $x = y$ and $x = -y$, which is not closed under addition. □

Problem 3. There is a 2×5 matrix A, s.t. $dim(Ker(A)) = 2$

Solution. False. Suppose it exists. From $dim(Ker(A)) = 2$ by rank-nullity theorem we have $rk(A) = 3$. But matrix with two rows cannot have rank more then two. This is contradiction, so such matrix cannot exist. □

Problem 4. a) There are 3×3 matrices A and B of rank 2, s.t. AB has rank 1.
b) There are 3×3 matrices A and B of rank 2, s.t. AB has rank 0.

Solution. a) True. Just take matrices which represent projections on x-y and y-x plane.
b) False. Suppose they exist. Then it means that AB is a matrix of zeroes, because it is the only matrix, whose rank is 0. This means that if $B = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix}$, then $AB = \begin{bmatrix} | & | & | \\ A\vec{v}_1 & A\vec{v}_2 & A\vec{v}_3 \\ | & | & | \end{bmatrix}$, and so $A\vec{v}_1 = A\vec{v}_2 = A\vec{v}_3 = \vec{0}$. Thus one has

$$Im(B) \subset Ker(A)$$

We know that rk(B)=dim(Im(B))=2, and rk(A)=2. Thus by rank-nullity dim(Ker(A))=1. Thus from $Im(B) \subset Ker(A)$ we get that 2 dimensional space is a subspace of 1 dimensional space — that is impossible, and so we have a contradiction. Thus such matrices do not exist. □

Problem 5. If A is 4×2 matrix, and B is 2×4 matrix, then $nullity(AB) \geqslant 2$.

Solution. True. By rank-nullity $nullity(AB) \geqslant 2 \iff rk(AB) \leqslant 2$. And this follows from rank inequality $rk(AB) \leqslant rk(A) \leqslant 2$, because A has 2 columns. □

Problem 6. $V^2, W^2 \subset \mathbb{R}^4$ are 2d subspaces, s.t. $V \cap W = \{\vec{0}\}$. Suppose $\{\vec{v}_1, \vec{v}_2\}$ is a basis of V, and $\{\vec{w}_1, \vec{w}_2\}$ is a basis of W. Then $\{\vec{v}_1, \vec{v}_2, \vec{w}_1, \vec{w}_2\}$ is a basis of $\mathbb{R}^4$

Solution. True. Suppose $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{w}_1 + c_4\vec{w}_1 = \vec{0}$. Then $c_1\vec{v}_1 + c_2\vec{v}_2 = -c_3\vec{w}_1 - c_4\vec{w}_1$, and because $V \cap W = \{\vec{0}\}$, we get $c_1\vec{v}_1 + c_2\vec{v}_2 = -c_3\vec{w}_1 - c_4\vec{w}_1 = \vec{0}$. Because $\{\vec{w}_1, \vec{w}_2\}$ and $\{\vec{v}_1, \vec{v}_2\}$ are bases, we get that $c_1 = c_2 = c_3 = c_4 = 0$, q.e.d. □

Problem 7. Matrices A and B have the largest possible rank given their size (min(n,m) for $n \times m$ matrix). Then their product AB also has the largest possible rank given its size.

Solution. False. Take $A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $B = [1, 0]$, then $AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ □

Problem 8. Suppose A is 2×4, and B is 4×7 matrices. Then

a) set of vectors in $\mathbb{R}^4$ which are simultaneously in Ker(A) and Im(B) is a subspace.

b) set of vectors in $\mathbb{R}^4$ which are either in Ker(A) or Im(B) is a subspace.

Solution. a) True. Intersection of subspaces is always a subspace.

b) False. Union of subspace is not always a subspace. □

Problem 9. Suppose A is 3×2, and B is 2×3 matrices. Then it can happen that $AB = I_3$.

Solution. False. Rank inequality. □

Review and True/False questions, Bretscher 3.4,5.1

Review

3.4 Change of coordinates

First of all, recall that our linear spaces $\mathbb{R}^n$ always come with a distinguished standard basis

$$\left\{ e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\}$$

Recall that $\mathbb{R}^n$ is a space of length n sequences $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ of numbers, which we call vectors. For every vector one has $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$, and thus one calls $x_1, \dots, x_n$ coordinates of vector $\vec{x}$ w.r.t. basis $\{\vec{e}_1, \dots, \vec{e}_n\}$. The main question of this section is "What happens if one would a pick a different basis of $\mathbb{R}^n$"?

Suppose $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a different (from standard) basis of $\mathbb{R}^n$. This, by definition, means that for every vector $\vec{x} \in \mathbb{R}^n$ we have a unique decomposition along the basis

$$\vec{x} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$$

We call numbers $c_1, \dots, c_n$ coordinates of vector $\vec{x}$ w.r.t. basis $\mathcal{B}$, and denote these coordinates by $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$.

Remark. If one denotes standard basis by $\mathcal{E}$, then one actually has $[\vec{x}]_{\mathcal{E}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \vec{x}$. So when we write a column of vectors, we always secretly mean that these are coordinates w.r.t. standard basis.

Example 1. Take $\mathcal{B} = \{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\}$. What are the coordinates $\begin{bmatrix}1\\1\end{bmatrix}_{\mathcal{B}}$, i.e. what are the coordinates of vector $\begin{bmatrix}1\\1\end{bmatrix}$ w.r.t. to a new basis ? What are the coordinates $\begin{bmatrix}-4\\0\end{bmatrix}_{\mathcal{B}}$? What are the coordinates $\begin{bmatrix}-2\\-1\end{bmatrix}_{\mathcal{B}}$?

Answer these questions by solving a linear system $\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2$. Illustrate the answers geometrically on the plane.

Main formulas

1. Change of coordinates formula. Suppose $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis of $\mathbb{R}^n$, and $S = \begin{bmatrix} | & & | \\ \vec{v}_1 & \dots & \vec{v}_n \\ | & & | \end{bmatrix}$. Then one has a formula, which relates coordinates of vectors w.r.t. standard basis $\{\vec{e}_1, \dots, \vec{e}_n\}$ and coordinates of vectors w.r.t. new basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$:

$$S \cdot [\vec{x}]_{\mathcal{B}} = \vec{x}$$

Because of this formula we sometimes call such S change of basis (or change of coordinates) matrix.

Example 2. For the previous example change of basis matrix would be $S = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$

2. Fix a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$. So far we always had a matrix

$$A = \begin{bmatrix} | & & | \\ Ae_1 & \dots & Ae_n \\ | & & | \end{bmatrix}$$

which represents transformation T. Turns out that this matrix represents T *with respect to standard basis*. What does it mean for matrix A' to represent T w.r.t. different basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$? This matrix should send $\mathcal{B}$-coordinates to $\mathcal{B}$-coordinates, i.e.: $A'[\vec{x}]_{\mathcal{B}} = [T(\vec{x})]_{\mathcal{B}}$. Thus, by definition, this matrix equals

$$A' = \begin{bmatrix} | & & | \\ [T(\vec{v}_1)]_{\mathcal{B}} & \dots & [T(\vec{v}_n)]_{\mathcal{B}} \\ | & & | \end{bmatrix}$$

(in analogy with matrix A, where $Ae_i = [T(e_i)]_{\mathcal{E}}$). Turns out that one can compute the matrix A' from matrices A and change of basis matrix S in this way:

$$A' = S^{-1}AS$$

Example 3. Understand what are A and A' matrices for the previous example, if T is a reflection about $x = y$ line. First spend some time to get the answer geometrically, from the definition of A', and then do a reality check by checking $A' = S^{-1}AS$.

These are equivalent definitions:

Definition 1. Two $n \times n$ matrices A and A' are called similar, if $A' = S^{-1}AS$ for some invertible matrix S.

Definition 2. Two $n \times n$ matrices A and A' are called similar, if there is such a linear transformation T and two bases $\mathcal{C}$ and $\mathcal{B}$ (one of them can be standard), s.t. A represents T w.r.t. $\mathcal{C}$, and A' represents T w.r.t. $\mathcal{B}$.

Similarity is transitive equivalence, i.e. if A is similar to B, B is similar to C, then A is similar to C. Also notice that A is similar to itself (take S=I), and notice that if A is similar to B, then B is similar to A (take S to be S^{-1}).

Another important fact is that similarity preserves rank.

5.1 Orthogonal projections and orthonormal bases

Orthogonality

Definition 3. Angle between two vectors is defined up to sign by

$$angle(\vec{v}, \vec{w}) = \frac{\vec{v} \cdot \vec{w}}{|\vec{v}| \cdot |\vec{w}|},$$

where at the top one has dot product, and in the bottom one has a product of lengths, where length is defined as usual $|\vec{x}| = \sqrt{x \cdot x} = \sqrt{x_1^2 + \cdots + x_n^2}$.

Definition 4. $\vec{v}$ is *perpendicular* to $\vec{w}$ if $\vec{v} \cdot \vec{w} = 0$. Notation $\vec{v} \perp \vec{w}$.

Definition 5. *Orthogonal compliment* of subspace $V \subset \mathbb{R}^n$ is a set $V^\perp$ of vectors which are perpendicular to every vector in V.

Proposition.

1. *$V^\perp$ is also a subspace of $\mathbb{R}^n$ (in the first place we had that V is a subspace).*
2. *$dim(V^\perp) = n - dim(V)$*
3. *$(V^\perp)^\perp = V$*
4. *$V \cap V^\perp = \vec{0}$*

Example 4. What is orthogonal compliment of line $x = y$ on the plane?

Definition 6. Vectors $\vec{v}_1, \ldots, \vec{v}_k$ are called *orthonormal* if they are pairwise orthogonal length 1 vectors. In other words:

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{otherwise} \end{cases}$$

Fact. Orthonormal vectors are always linearly independent. Thus any n orthonormal vectors form a basis of $\mathbb{R}^n$.

We call such bases *orthonormal basis.* They behave like standard basis, and are very important. One can think of such basis as the image of standard basis after applying rotation or reflection.

Any subspace $V \subset \mathbb{R}^n$ has an orthonormal basis (not unique), and later we will see how to find one.

Orthogonal projection

If V is a subspace of $\mathbb{R}^n$, then any $\vec{x} \in \mathbb{R}^n$ can be split into sum

$$\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$$

where $\vec{x}^{\parallel} \in V$ is the parallel to V part, and $\vec{x}^{\perp} \in V^{\perp}$ is the perpendicular to V part. Notice, that since $\vec{x}, \vec{x}^{\parallel}, \vec{x}^{\perp}$ form a triangle with the right angle, one has inequalities of length: $|\vec{x}^{\perp}| \leqslant |\vec{x}|, |\vec{x}^{\parallel}| \leqslant |\vec{x}|$.

Definition 7. *Orthogonal projection* of vector $\vec{x} \in \mathbb{R}^n$ onto subspace V is the parallel part of the vector: $proj_V \vec{x} = \vec{x}^{\parallel}$.

Transformation $T(\vec{x}) = proj_V \vec{x}$ is linear, and thus has its matrix. We will have a formula

Formula for orthogonal projection. If $\{\vec{v}_1, \dots, \vec{v}_k\}$ is orthonormal basis of $V \subset \mathbb{R}^n$, then

$$proj_V \vec{x} = (\vec{x} \cdot \vec{v}_1)\vec{v}_1 + \cdots + (\vec{x} \cdot \vec{v}_k)\vec{v}_k$$

If one is interested in the matrix P, which represent orthogonal projection, then the best way would be to compute images of standard basis $T(\vec{e}_i) = proj_V \vec{e}_i$ based on the formula above, and then use the fact that $P = \begin{bmatrix} | & & | \\ T(\vec{e}_1) & \dots & T(\vec{e}_n) \\ | & & | \end{bmatrix}$

Remark. Orthogonal projection matrix P (on a subspace V) is similar to a matrix with 1's and zeroes on the diagonal, of the following form: (number of ones is equal to rank(P)=dim(V)).=
$$\begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix}$$
Why is this true?

True/False questions

The following problem is a typical example of true/false question on similarity of matrices.

Problem 1. Rotations by $\pi/2$ and $\pi/4$ are similar.

Solution. False. If they are similar, then $SR_{\pi/2} = R_{\pi/4}S$ for some invertible 2×2 matrix S. Suppose

$$S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Then, multiplying with rotation matrices, one gets

$$\begin{bmatrix} b & -a \\ d & -c \end{bmatrix} = 1/\sqrt{2} \begin{bmatrix} a-c & b-d \\ c+c & d+d \end{bmatrix}$$

From this it is an exercise to prove that S has to be a zero matrix, which is not invertible. This is contradiction, and thus two rotations cannot be similar. □

Problem 2. If A and B are invertible $n \times n$ matrices, then AB is similar to BA

Solution. True, take $S = B$. □

Problem 3. If A and B are similar, then A+I and B+I are similar

Solution. True, if SA=BS, then S(A+I)=(B+I)S. □

sad

Problem 4. If F is a non-zero 2×2 matrix, s.t. $F\vec{x} = 0$ for every $\vec{v} \perp \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then F is a matrix of orthogonal projection.

Solution. False, take $F = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ □

Problem 5. If $A^2 = A$, then matrix A is an orthogonal projection onto a subspace of $\mathbb{R}^n$.

Solution. False, take $A = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ □

Problem 6. Matrices $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ are similar

Solution. True, right down equation SA=BS, solve for $S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, and find out that there is a solution where S is invertible. □

Problem 7. There are four orthonormal vectors in $\mathbb{R}^3$.

Solution. False, 4 vectors cannot be linearly independent in 3D. □

Problem 8. There exist four vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ in $\mathbb{R}^2$, s.t. $\vec{v}_i \cdot \vec{v}_j < 0$ for $i \neq j$.

Solution. False, because angles between them should be more then $90°$. □

Problem 9. There exist four vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ in $\mathbb{R}^2$, s.t. $\vec{v}_i \cdot \vec{v}_j < 0$ for $i \neq j$.

Solution. False, because angles between them should be more then $90°$. □

Problem 10. $T = \begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}$ is similar to a matrix of some orthogonal projection.

Solution. True. Solve $ST = PT$, where $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $S = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. (we picked P in that form because we know that T has rank 1, and so if there is similar projection matrix, it will have rank 1, and thus will be similar to P) □

Review and True/False questions, Bretscher 5.2-5.4

Review

5.2 Gram-Schmidt process, QR factorization

Gram-Schmidt process

Remark. We will do everything in the case where $V \subset \mathbb{R}^n$ is three-dimensional. Generalization to k-dimensional subspaces is straightforward.

Suppose $\{\vec{w}_1, \vec{w}_2, \vec{w}_3\}$ is a basis of 3D subspace $V^3 \subset \mathbb{R}^n$. Gram-Schmidt process is an algorithm, which allows one to change vector by vector in the basis, s.t. in the end one gets an orthonormal basis $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ of the same subspace $V^3 \subset \mathbb{R}^n$. The formulas are the following: (accompanied with geometric illustration on the blackboard)

$$\begin{cases} \vec{u}_1 = & \frac{\vec{w}_1}{|\vec{w}_1|} \\ \vec{u}_2 = & \frac{\vec{w}_2^\perp}{|\vec{w}_2^\perp|}, \text{ where } \vec{w}_2^\perp = \vec{w}_2 - (\vec{u}_1 \cdot \vec{w}_2)\vec{u}_1 \\ \vec{u}_3 = & \frac{\vec{w}_3^\perp}{|\vec{w}_3^\perp|}, \text{ where } \vec{w}_3^\perp = \vec{w}_3 - (\vec{u}_1 \cdot \vec{w}_3)\vec{u}_1 - (\vec{u}_2 \cdot \vec{w}_3)\vec{u}_2 \end{cases}$$

Notice, that one has to compute $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ in the order, because formula for $\vec{u}_2$ involves $\vec{u}_1$, and formula for $\vec{u}_3$ involves $\vec{u}_1, \vec{u}_2$.

QR factorization

Suppose we are given a matrix $A_{l\times k}$, whose columns are linearly independent (i.e. columns form basis of Im(A)). Then, using Gram-Schmidt process, one can obtain *QR factorization*

$$A_{l\times k} = Q_{l\times k} R_{k\times k}$$

where Q is a matrix whose columns form orthonormal basis of Im(A), and R is upper triangular matrix.

Columns of Q can be found by running G-S process on columns of A. For k=3 the formula for factorization is the following:

$$A = \begin{bmatrix} | & | & | \\ \vec{w}_1 & \vec{w}_2 & \vec{w}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \\ | & | & | \end{bmatrix} \cdot \begin{bmatrix} |\vec{w}_1| & \vec{u}_1 \cdot \vec{w}_2 & \vec{u}_1 \cdot \vec{w}_3 \\ 0 & |\vec{w}_2^\perp| & \vec{u}_2 \cdot \vec{w}_3 \\ 0 & 0 & |\vec{w}_3^\perp| \end{bmatrix} = QR$$

Suppose A is $n \times n$ square matrix of full rk(A)=n. Then in factorization $A = QR$ the Q matrix is orthogonal.

Formula for the projection matrix. Suppose $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$ is a basis of a subspace $V \subset \mathbb{R}^n$. Then if $A = \begin{bmatrix} | & & | \\ \vec{w}_1 & \ldots & \vec{w}_k \\ | & & | \end{bmatrix}$, one has $V = Im(A)$. Suppose $A = QR$ is QR-factorization of A. Then the matrix $Q \cdot Q^T$ is a matrix of orthogonal projection on V=Im(A). In other words:

$$proj_{Im(A)}\vec{x} = QQ^T \cdot \vec{x}$$

Remark. One very useful (and easy to check) property of orthogonal projection matrices is that $P^2 = P$.

Definition 1. Suppose matrix $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \ldots & \vec{v}_k \\ | & & | \end{bmatrix}$. Then transpose of it is a matrix $A^T = \begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_k & - \end{bmatrix}$. In other words transposing corresponds to reflection of the matrix about it's $-45°$diagonal (notice, that matrix doesn't have to be square).

Properties of transpose matrices.

1. $(A^T)^T = A$
2. $(AB)^T = B^T A^T$
3. $(A^{-1})^T = (A^T)^{-1}$
4. $rk(A^T) = rk(A)$
5. $Ker(A)$ is orthogonal compliment of $Im(A^T)$
6. (dot product) $\vec{v} \cdot \vec{w} = \vec{v}^T \cdot \vec{w} = \vec{w}^T \cdot \vec{v}$ (products of matrices)

Warning. $rref(A)^T \neq rref(A^T)$

Definition 2. Symmetric matrices are those, which satisfy $A = A^T$. (makes sense only for square matrices)

Example 1. Orthogonal projection matrices P are symmetric, because

$$P^T = (QQ^T)^T = (Q^T)^T Q^T = QQ^T = P$$

5.3 Orthogonal matrices

Definition 3. The following are four equivalent definitions of *orthogonal* square $n \times n$ matrix A:

1. $(A^T)A = I_n = A(A^T)$
2. $A^{-1} = A$

3. columns of A form orthonormal basis of $\mathbb{R}^n$

4. Matrix A preserves dot product, i.e. $\vec{x} \cdot \vec{y} = A\vec{x} \cdot \vec{A}y$ for any two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$. This is equivalent to saying that A preserves lengths, i.e. $|A\vec{x}| = |\vec{x}|$ for any $\vec{x} \in \mathbb{R}^n$

Remark. The terminology is a bit confusing, so it is good to have a clear distinction in mind between "orthogonal projection matrices", which never have a full rank, and "orthogonal matrices", which always have a full rank.

Example 2. What are orthogonal 2×2 matrices? Turns out those are all rotation matrices, and rotation composed with reflection matrices.

Properties of orthogonal matrices.

1. Orthogonal matrices always have full rank.
2. A is orthogonal $\implies$ A^T is orthogonal
3. A is orthogonal $\implies$ $det(A) = \pm 1$
4. $rk(A^T) = rk(A)$
5. Orthogonal projection matrices are never orthogonal matrices

5.4 Least squares method

Suppose we are given a system $A\vec{x} = \vec{b}$ which doesn't have a solution. It means that the subspace $Im(A)$ misses the vector $\vec{b}$ (illustration on the blackboard).

Then the best thing one would hope is to find such $\vec{b}^* \in Im(A)$, that $dist(\vec{b}^*, \vec{b})$ is minimal, and then find its preimage $\vec{x}^*$. This $\vec{x}^*$ is called least squares solution. This is equivalent to condition

$$A\vec{x}^* = proj_{Im(A)}\vec{b} \quad (= \vec{b}^*)$$

Turns out that there is a very convenient way to find $\vec{x}^*$, namely one has to solve

$$A^T A = A^T \vec{b}$$

Example 3. Fitting quadric to four points on the plane (-1.8), (0,8), (1,4), (2,16) — example 4 from the book.

Example 4 (T/F question)**.** We are looking for least squares solution to $A_{4\times 3}\vec{x} = \vec{b}$. We know that $Im(A)^{\perp} = span(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix})$, $\vec{b} = span(\begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \end{bmatrix})$. Then the claim is that $A\vec{x}^* = \begin{bmatrix} 1 \\ 1 \\ 2 \\ -2 \end{bmatrix}$

Solution. False. Let's prove it by contradiction: suppose its true. Key observation is that because $A\vec{x}^* = proj_{Im(A)}\vec{b}$ one has $\vec{b} - A\vec{x}^* \perp Im(A)$. Thus one has to have $\vec{b} - A\vec{x}^* = k \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, which is not true in this case — contradiction. $\square$

True/False questions

Problem 1. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear transformation. Then if $m \neq n$, then T cannot preserve lengths.

Solution. False. Counterexample is an embedding of plane into 3D space, given by matrix
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$
□

Problem 2. Matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ represents orthogonal projection.

Solution. False. 1st way to see it: $A^2 \neq A$. 2nd way: the matrix has full rank. □

Problem 3. If A is 2×2 matrix, such that A^2 is orthogonal, then A is also orthogonal.

Solution. False. Counterexample is $\begin{bmatrix} 0 & 2 \\ 1/2 & 0 \end{bmatrix}$. $A^2 = I_2$ is orthogonal, whereas A is not, because doesn't preserve lengths.

How one would find this solution? A possible approach is to just take matrix A in the general form $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then write down what does it mean to have A^2 orthogonal, in those equations assume $a = -d$, and the solution will come out. □

Problem 4. If B is orthogona, A is similar to B, then A is orthogonal.

Solution. False. Take A to be matrix of linear transformation which swaps vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (what is this matrix?). Then we have $S^{-1}AS = B$, where $S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (B is a matrix representing the same linear transformation as A, but in basis of columns of S). Thus B is similar to A. B is clearly orthogonal, but A is not, because it doesn't preserve lengths (because it swaps $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$). Thus we found counterexample. □

Problem 5. Suppose K is 2×2 a matrix of a projection on x-axis, L is a matrix of a projection on y-axis 2×2. Then KL, $(KL)^2$, $(KL)^3$, ... are all different matrices.

Solution. False. Compute K and L by the formula for orthogonal projection and notice that all KL, $(KL)^2$, $(KL)^3$, ... are zero matrices. □

Problem 6. A is similar to $A^T \implies$ A is symmetric.

Solution. False. Write down general equation for 2×2 case $SA = A^T S$, and see that there is a lot of freedom in choosing coefficients of S and A, allowing A being not symmetric. □

Problem 7. Suppose $\{\vec{v}_1, \ldots, \vec{v}_6\}$ is a basis of $\mathbb{R}^6$. By Gram-Schmidt one gets orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_6\}$. Suppose $V = span(\vec{v}_1, \ldots, \vec{v}_5)$, $\vec{x} = proj_V \vec{u}_6$. Then it is possible that $|\vec{x}| \geqslant |\vec{v}_1|/\sqrt{5}$

Solution. False. From the way G-S process works one gets $V = span(\vec{v}_1, \ldots, \vec{v}_5) = span(\vec{u}_1, \ldots, \vec{u}_5)$. Thus $\vec{u}_6 \perp V$. Thus $\vec{x} = proj_V \vec{u}_6 = \vec{0}$, and its length is of course 0. □

Problem 8. There is 3 matrix such that $Ker(A) = Im(A)$

Solution. False. From rank-nullity theorem one gets that $dim(Ker(A)) = dim(Im(A)) = 3/2$, which is not possible. □

Review and True/False questions, Bretscher 6.1-6.3

Review

6.1 Definitions of determinant

1×1 **case** $\det([a]) = a$.

2×2 **case** $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.

3×3 **case** $\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} = aek + bfg + dhc - gec - dbk - afh$.

Heuristics to remember the formula: Sarrus's rule (p.266 of the book), or via patterns (see below).

$n \times n$ **case** Suppose A is $n \times n$ matrix. Then

$$\det A = \sum_{Patterns\ P} (-1)^{inv(P)} \cdot Prod(P)$$

where pattern of matrix $A_{n \times n}$ is a set of n coefficients of matrix such that each row and column get one, like here for example: pattern $S = \begin{bmatrix} \cdot & \circ & \cdot & \cdot \\ \cdot & \cdot & \circ & \cdot \\ \circ & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \circ \end{bmatrix}$.

By $Prod(P)$ we denote the product of all elements of the pattern, and by $inv(P)$ we denote the number of inversions, i.e. all the pairs (a, b) of elements in P such that line going through a and b on the matrix A has positive slope. For example $inv(S) = 2$.

Reality check. Check that definition for $n \times n$ matrix generalizes three previous definitions.

Example 1. Compute $\det \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$

Big questions for the next section are: what are determinants useful for, and how do we compute them?

6.2 Properties of determinant

1. $\det A = \det A^T$ (this implies that any property about columns below has analogous property for rows).

2. Linearity w.r.t. rows and columns:

$$\det \begin{bmatrix} | & & | & & | \\ \vec{v}_1 & \dots & a\vec{x}+b\vec{y} & \dots & \vec{v}_n \\ | & & | & & | \end{bmatrix} = a{\cdot}\det \begin{bmatrix} | & & | & & | \\ \vec{v}_1 & \dots & \vec{x} & \dots & \vec{v}_n \\ | & & | & & | \end{bmatrix} + b{\cdot}\det \begin{bmatrix} | & & | & & | \\ \vec{v}_1 & \dots & \vec{y} & \dots & \vec{v}_n \\ | & & | & & | \end{bmatrix}$$

3. Swapping two rows (columns) of matrix A changes the sign of the $\det A$.

4. A has identical columns (rows) $\implies \det A = 0$.

5. $\det AB = \det A \cdot \det B$.

6. This theorem is the main reason for introducing the concept of determinant:

   Theorem. *Square matrix A is invertible (i.e. has maximal rank, zero kernel, its reduced row echelon form is identity, etc) if and only if its determinant is not zero.*

How to compute determinants

If size of the square matrix is $\leqslant 3$, then the easiest thing is to just use the formula. In general though, for $n \times n$ matrix one has $n!$ number of patterns. For $n = 4$ one has $4! = 4 * 3 * 2 * 1 = 24$ patterns, which is already unpleasant to write down and compute their products. Thus one usually uses one of the following tricks:

Trick 1 (picking only non-zero patterns)**.** Sometimes, if matrix A has a lot of zeroes, there are only few pattern which do not have zero in it. In this case only these few patterns contribute to determinant.

Example 2.

$$\det \begin{bmatrix} 6 & 0 & 1 & 0 & 0 \\ 9 & 3 & 2 & 3 & 7 \\ 8 & 0 & 3 & 2 & 9 \\ 0 & 0 & 4 & 0 & 0 \\ 5 & 0 & 5 & 0 & 4 \end{bmatrix} = (-1)^1 \cdot 6 \cdot 3 \cdot 2 \cdot 4 \cdot 4 = -288.$$

Example 3. Determinant of any upper-triangular or lower-triangular square matrix equals to the produc of diagonal elements.

Trick 2 (simplifying matrix using elementary row (column) operations)**.** By the properties of determinant we know how det changes with elementary row (column) operations:

- Multiplying row by k results into multiplying determinant by k (property 2).
- Swapping two rows changes the sign of determinant (property 3).

- Adding a row with coefficient to another row preserves the determinant (check this using properties 2 and 4).

Thus, even if one has a big matrix A with a lot of non-zero elements, one can reduce it to $rref(A)$ (where there are lots of zeroes) and then use the above properties to compute the determinant of the initial matrix A. This is Algorithm 6.2.5 on p.282 of the book.

Notice, that sometimes its not necessary to do it all the way to the $rref(A)$, as one can guess the determinant by Trick 1 on half of the way.

Example 4.

$$\det \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \det \begin{bmatrix} -1 & -1 & -1 & -1 & 5 \\ 0 & -1 & -1 & -1 & 4 \\ 0 & 0 & -1 & -1 & 3 \\ 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} = (-1) \cdot (-1) \cdot (-1) \cdot (-1) \cdot 1 = 1$$

The first equality we have because one gets the second matrix from the first by subtracting second column from first, third column from second, forth from third, and fifth from forth. The second equality is done remembering what is determinant for upper-triangular matrices.

Trick 3 (Laplace expansion along the column (row))**.** Suppose one does not have a lot of zeroes in a matrix A, but one row (column) does have a few zeros. Then Laplace expansion is a very useful tool to compute determinants. See Theorem 6.2.10 in the book, and example after it.

6.3 Geometric interpretation of determinant

Suppose linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is represented by square matrix A. Then, if $\Omega \subset \mathbb{R}^n$ is an open region, one has the following

Theorem.

$$|\det(A)| = \text{"}expansion\ factor\text{"} = \frac{Volume(A(\Omega))}{Volume(\Omega)}$$

A special case of this is the following. A parallelepiped generated by $(\vec{v}_1, \ldots, \vec{v}_n)$ is the following set

$$P(\vec{v}_1, \ldots, \vec{v}_n) = \{a_1\vec{v}_1 + \cdots + a_n\vec{v}_n |\ 0 < a_i < 1\}$$

Volume of unit cube is 1, so one has the following

Proposition. *For matrix* $A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \ldots & \vec{v}_n \\ | & & | \end{bmatrix}$ *one has* $P(\vec{v}_1, \ldots, \vec{v}_n) = |\det A|$.

Example 5 (2×2 case)**.** The area of parallelogram $P(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix})$ is equal to $\det A = |x_1 \cdot y_2 - x_2 \cdot y_1|$.

True/False questions

Problem 1. $\det(4 \cdot A) = 4 \cdot \det A$ for 3×3 matrices. What about 1×1 matrices?

Solution. For 3×3 false, for 1×1 true. Use linearity w.r.t. columns:

$$\det \begin{bmatrix} | & | & | \\ 4\vec{v}_1 & 4\vec{v}_2 & 4\vec{v}_3 \\ | & | & | \end{bmatrix} = 4^3 \cdot \det \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix}$$

□

Problem 2. $\det(A + B) = \det A + \det B$

Solution. False. Take, for example, two identity matrices. □

Problem 3. $A = \begin{bmatrix} k^2 & 1 & 4 \\ k & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$ is invertible for all k.

Solution. $\det A = -3k^2 + 3k + 6 = 0$ for $k = -1$, and so A is not always invertible by Theorem 6. □

Problem 4. There is a 3×3 matrix A s.t. $A^2 = -I_3$.

Solution. False. Otherwise one has a real number which squares to -1: $(\det A)^2 = \det A^2 = \det(-I_3) = -1$, and this is a contradiction. □

Problem 5. A is 3×2 and B is 2×3. Then $\det(AB) = \det(BA)$.

Solution. False. One would guess this by the fact that AB is a 3x3 matrix whose rank is not full (by $rk(AB) \leqslant rk(A) \leqslant 2$), and thus $\det(AB) = 0$, whereas $\det(BA)$ should not be 0 necessarily, because it is 2x2.

As a counterexample one can take $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ (these are first goto non-square matrices one usually checks) □

Problem 6. A is orthogonal $\implies$ $\det A = \pm 1$.

Solution. True. From definition of orthogonal matrices on has $A \cdot A^T = I_n$, and by taking determinant of both sides one gets the desired statement (remembering properties 1 and 5 of determinant). □

Problem 7. $\det A = \pm 1 \implies$ A is orthogonal.

Solution. False. Take for example matrix A which stretches by 2 in horizontal direction, and shrinks by 1/2 in vertical direction. $A = \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}$. □

Problem 8. $A = \begin{bmatrix} | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{v}_4 \\ | & | & | & | \end{bmatrix}$, $B = \begin{bmatrix} | & | & | & | \\ \vec{v}_1 + \vec{v}_2 + \vec{v}_3 & 3\vec{v}_1 + \vec{v}_2 & \vec{v}_3 + 4\vec{v}_4 & 5\vec{v}_4 \\ | & | & | & | \end{bmatrix}$. Then $|\det B| = 10|\det A|$.

Solution. True. Follows from the fact that $B = A \cdot \begin{bmatrix} 1 & 3 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 4 & 5 \end{bmatrix}$ (notice how columns of this matrix depend on expressions of columns of B in terms of columns of A).

One then has

$$\det B = \det A \cdot \det \begin{bmatrix} 1 & 3 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 4 & 5 \end{bmatrix} \implies (\textit{by trick } 1) \det B = \det A \cdot -10 \implies$$

$$\implies (\textit{take absolute value}) |\det B| = 10|\det A|$$

$\square$

Review and True/False questions, Bretscher 7.1-7.2

Review

7.1 Diagonalization

We call square matrices *diagonal* if their non-diagonal elements are all 0, i.e. matrix (for 3x3 case) is of the form

$$\begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{bmatrix}$$

(notice that c_i could be equal to 0)

Diagonal matrices are very nice: we understand exactly what they are doing as linear transformations (stretching by c_i along $i-th$ coordinate), and it is also very easy to multiply them, and in particular compute their powers:

$$\begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{bmatrix}^t = \begin{bmatrix} c_1^t & 0 & 0 \\ 0 & c_2^t & 0 \\ 0 & 0 & c_3^t \end{bmatrix}$$

Thus, having a matrix A with lots of zeroes, we want to "make" it diagonal, i.e. to diagonalize it. What does it mean? It means to find a basis, in which matrix A becomes diagonal. Let us give two precise definitions of it:

Definition 1. To *diagonalize* square matrix $A_{n\times n}$ means to find a diagonal matrix B, which is similar to A. I.e. for 3x3 case for example, it means that

$$A = SBS^{-1} = S \cdot \begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{bmatrix} \cdot S^{-1}$$

for some invertible change of basis matrix $S = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix}$.

Let us say the same thing differently.

Definition 2. Suppose T is linear transformation corresponding to matrix $A_{3\times3}$. Diagonalizing matrix A means to find a basis $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$, w.r.t. which linear transformation T has a diagonal matrix $B = \begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{bmatrix}$. This is equivalent to finding a basis $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$

such that $A\vec{v}_i = c_i \cdot \vec{v}_i$. We call such basis *eigenbasis*, the vectors $\vec{v}_i$ are *eigenvectors*, and numbers c_i are corresponding *eigenvalues* of eigenvectors $\vec{v}_i$. It is important to notice, that eigenvectors are far from being unique, there is a whole vector space of eigenvectors. The way we capture that space is by picking its basis.

Remark. Notice, that after finding eigenvectors and eigenvalues we basically know everything, because matrix B is diagonal with eigenvalues on its diagonal, and change of basis matrix S have eigenvectors as its columns.

Two above definitions are equivalent because of the way matrices change, when one changes the basis: for change of basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3) \to (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ the matrix changes according to the rule $A = SBS^{-1}$, where $S = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix}$ is a change of basis matrix.

Example 1 (with a picture on the blackboard)**.** Let us diagonalize matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. For this, let us think what would be an eigenbasis for such matrix. From the columns of this matrix it follows that it swaps the standard basis vectors: $A\vec{e}_1 = \vec{e}_2,\ A\vec{e}_2 = \vec{e}_1$. Thus this matrix must represent a reflection about the line $x = y$. Thus, a natural eigenbasis would be $(\vec{e}_1 + \vec{e}_2, \vec{e}_1 - \vec{e}_2)$, where first vector is preserved, and the second vector is multiplied by (-1). Thus we must have

$$A = SBS^{-1} \iff \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}$$

Warning. Not every matrix is diagonalizable. The main two examples are rotation matrix, say $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, and the other canonical example is a shear matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, see below for why they are not diagonalizable.

7.2 Eigenvalues, characteristic polynomial

This and the next section will answer the question "How do we diagonalize matrix A, and what can go wrong?".

Algorithm (How to diagonalize matrices?)**.** Having a square matrix $A_{n\times n}$ the step by step procedure for diagonalization is the following:

1. Find all eigenvalues $\lambda_1, \ldots, \lambda_n$. See below for how to do it and what can go wrong.

2. If one found exactly n eigenvalues (counted with multiplicities), then one proceeds to find linearly independent eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$. How to do it and what can go wrong is the next section 7.2.

3. If one was able to get exactly n linearly independent eigenvectors in the end, then $(\vec{v}_1, \ldots, \vec{v}_n)$ is an eigenbasis and the diagonalization is finished. Let us repeat what it means. Suppose A represents linear transformation T. Then w.r.t. eigenbasis

$(\vec{v}_1, \ldots, \vec{v}_n)$ T has diagonal matrix $B = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$. Said differently, one has the following formula representing the fact that matrix B is *similar* to matrix A:

$$A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \ldots & \vec{v}_n \\ | & & | \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \cdot \begin{bmatrix} | & & | \\ \vec{v}_1 & \ldots & \vec{v}_n \\ | & & | \end{bmatrix}^{-1}$$

How to find eigenvalues

The way to find eigenvalues it is to use the following

Theorem (Eigenvalues are roots of char polynomial)**.** *Eigenvalues of matrix $A_{n\times n}$ are precisely the roots of characteristic polynomial in t:*

$$f_A(t) = \det(A - t \cdot I_n)$$

Proof. λ is eigenvalue of A $\iff$ there is such a vector $\vec{v} \neq 0$ that $A\vec{v} = \lambda \cdot \vec{v} \iff A\vec{v} - \lambda \cdot \vec{v} = 0 \iff A\vec{v} - \lambda I_n \vec{v} = 0 \iff (A - \lambda I_n)\vec{v} = 0 \iff A - \lambda I_n$ has non-zero kernel $\iff A - \lambda I_n$ is not invertible $\iff \det(A - \lambda I_n) = 0 \iff \lambda$ is a root of $f_A(t)$. $\square$

Thus to find eigenvalues of A, just compute its characteristic polynomial $f_A(t)$, and then find its roots. In case, where one has roots with different multiplicities in the polynomial, say $f_A(t) = (t-4)^2(t+7)^6$, then one says that eigenvalues $\lambda = 4$ has *algebraic multiplicity* 2, and eigenvalue $\lambda = -7$ has its *algebraic multiplicity* 6. May be even better way go about, is to just think that there are still 8 eigenvalues, but some of the are equal, namely $\lambda_1 = \lambda_2 = 4$ and $\lambda_3 = \cdots = \lambda_8 = -7$.

Definition 3. *Algebraic multiplicity* of eigenvalue λ of matrix $A_{n\times n}$ is the maximal degree $k = almu(\lambda)$ such that $(\lambda - t)^k g(t) = f_A(t)$

What can go wrong. The problem occurs when we are not able to find exactly n roots, counted with multiplicities. This happens only if one cannot decompose polynomial $f_A(t)$ into linear terms. This happens only if $f_A(t) = (t^2+pt+q)g(t)$, and the quadratic polynomial has negative discriminant.

Example 2 (R_{90^o} does not have real eigenvalues)**.** If one takes a rotation matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then one has

$$\det(A - tI_2) = \det \begin{bmatrix} -t & -1 \\ 1 & -t \end{bmatrix} = t^2 + 1$$

which is irreducible over reals (its roots are imaginary). Thus, working over reals, this rotation matrix does not have eigenvalues.

Question. When does rotation matrix have real eigenvalues? Does this happen at all?

Remark. This problem occurs only if one restricts them-self to the field of real numbers $\mathbb{R}$. The problem does not occur if one works over field $\mathbb{C}$, see section 7.5 for that.

Example 3 (Reality check for the example from the previous section)**.** We return to our example from the previous section. There, for $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ we just guessed eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, and then the eigenvalues are 1 and -1, because one eigenvector is preserved by A, and the other is flipped to its negative. Now, let us check, that ± 1 are indeed the roots of $f_A(t)$:

$$\det(A - tI_2) = \det \begin{bmatrix} -t & 1 \\ 1 & -t \end{bmatrix} = t^2 - 1 = (t-1)(t+1)$$

and indeed ± 1 are the roots of this polynomial.

Discrete dynamical systems, and how eigenvectors help there

Discrete dynamical system is described by (1) dynamical system equation (2) initial value. It is given by system

$$\begin{cases} \vec{x}(n+1) = A \cdot x(n) \\ \vec{x}(0) = \vec{x}_0 \end{cases}$$

All the questions about dynamical systems usually boil down to formula for $\vec{x}(n)$ in terms of matrix A and initial value $\vec{x}_0$. Here we treat the case where A is 2x2 diagonalizable matrix.

First, notice that from the system it follows that

$$\vec{x}(n) = A^n \vec{x}_0$$

thus it is would be enough to find a formula for A^n. This is the idea behind the first method:

First method (via computing A^n)**.** Suppose we diagonalized A, i.e. $A = S \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} S^{-1}$. Then it is easy to compute any of its powers:

$$A^n = (SBS^{-1})^n = SB^nS^{-1} = S \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} S^{-1}$$

and so

$$x(n) = S \begin{bmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{bmatrix} S^{-1} \cdot x_0$$

Overall it is a good method, but it can be difficult to compute S^{-1}. Thus, we think that the following is a slightly better approach

Second method (via decomposing x_0 along eigenvectors)**.** Again, suppose we diagonalized A, i.e. we found an eigenbasis $(\vec{v}_1, \vec{v}_2)$ with eigenvalues λ_1, λ_2. Then let us consider three cases.

1. Case $x_0 = \vec{v}_1$. Then from $A\vec{v}_1 = \lambda_1 \vec{v}_1$ it is clear that $A^n x_0 = A^n \vec{v}_1 = \lambda_1^n \vec{v}_1$.

2. Case $x_0 = \vec{v}_2$. Then from $A\vec{v}_2 = \lambda_2 \vec{v}_2$ it is clear that $A^n x_0 = A^n \vec{v}_2 = \lambda_2^n \vec{v}_2$.

3. General case $x_0 = c_1\vec{v}_1 + c_2\vec{v}_2$. Then by linearity one gets

$$A^n x_0 = A^n(c_1\vec{v}_1 + c_2\vec{v}_2) = c_1\lambda_1^n\vec{v}_1 + c_2\lambda_2^n\vec{v}_2$$

The main difficulty in this step is to find decomposition $x_0 = c_1\vec{v}_1 + c_2\vec{v}_2$ (which exists because $(\vec{v}_1, \vec{v}_2)$ is a basis). This can be done by solving for c_1, c_2 the linear system $x_0 = c_1\vec{v}_1 + c_2\vec{v}_2$. Secretly this computation is the same as computation of S^{-1} in the previous method.

True/False questions

We start from two typical problems on eigenvalues.

Problem 1. If $\vec{v}$ is an eigenvector of A, then it is an eigenvector for $A^3 + A^2 - A$.

Solution. True. $(A^3 + A^2 - A)\vec{v} = A^3\vec{v} + A^2\vec{v} - A\vec{v} = \lambda^3\vec{v} + \lambda^2\vec{v} - \lambda\vec{v} = (\lambda^3 + \lambda^2 - \lambda)\vec{v}$. □

Problem 2. 3x3 matrix satisfies $B^3 = 0$. Then the real eigenvalues of B can only be 0.

Solution. True. Because is λ is eigenvalue of B, then λ^3 is an eigenvalue of a zero matrix $B^3 = 0$, and thus $\lambda^3 = 0$ (because zero matrix has only one zero eigenvalue). □

Problem 3. Every 5x5 matrix has real eigenvalue.

Solution. True. The degree of $f_{A_{n\times n}}(t)$ is equal to n, and so characteristic polynomial of 5x5 matrix has degree 5. Every odd degree polynomial has a real root (by understanding what happens as $t \to \pm\infty$ and mean value theorem), thus 5x5 matrix must have a real eigenvalue. □

Problem 4. All orthogonal matrices can have only eigenvalues ± 1.

Solution. True. Because orthogonal matrices preserve lengths, and thus can multiply eigenvectors only by ± 1. □

Problem 5. A is 3x2 matrix, B is 2x3 $\implies$ 0 is an eigenvalue of AB.

Solution. True. First, from definition "0 being eigenvalue" is equivalent to $\det(AB) = 0$. Thus, we need to prove that AB is not invertible. But this follows from the fact that it doesn't have full rank:

$$rk(AB_{3\times 3}) \leqslant rk(A_{3\times 2}) \leqslant 2$$

□

Problem 6. A is 2x3 matrix, B is 3x2, $\lambda \neq 0$ is an eigenvalue of $BA \implies \lambda$ is an eigenvalue of AB.

Solution. True. $BA\vec{v} = \lambda\vec{v} \implies ABA\vec{v} = A\lambda\vec{v} \implies (AB)\cdot(A\vec{v}) = \lambda\cdot(A\vec{v})$, and $(A\vec{v})$ is a non-zero vector since we had $\lambda \neq 0$ and $A\vec{v} = \lambda\vec{v}$. □

Problem 7. For square matrices A^T and A have the same eigenvalues.

Solution. True. Because $\det A^T = \det A$, one has that if $\det(A - \lambda I_n) = 0$, then $\det((A - \lambda I_n)^T) = \det(A^T - \lambda I_n^T) = \det(A^T - \lambda I_n) = 0$. $\square$

Problem 8. 2x2 matrices D and E have eigenvalues $\lambda_1 = \lambda_2 = 1$ for D and $\lambda_1 = \lambda_2 = 3$ for E. Then D+E has eigenvalues 4,4.

Solution. False. Take $D = \begin{bmatrix} 1 & 8 \\ 0 & 0 \end{bmatrix}$ and $E = \begin{bmatrix} 3 & 0 \\ 8 & 3 \end{bmatrix}$. $\square$

Review and True/False questions, Bretscher 7.3,7.5

Review

7.3 Eigenvectors, eigenspaces

Recall, that we are discussing the following algorithm.

Algorithm (How to diagonalize matrices?)**.** Having a square matrix $A_{n\times n}$ the step by step procedure for diagonalization is the following:

1. Find all eigenvalues $\lambda_1, \dots, \lambda_n$ by finding roots of characteristic polynomial $f_A(t) = \det(A - t \cdot I_n)$ (possibly with multiplicities).

2. Suppose one found exactly n eigenvalues (counted with multiplicities). Then one proceeds to linearly independent eigenvectors $\vec{v}_1, \dots, \vec{v}_n$. Below we describe this process, and what can go wrong.

3. If one was able to get exactly n linearly independent eigenvectors in the end, then $(\vec{v}_1, \dots, \vec{v}_n)$ is an eigenbasis and the diagonalization is finished. Let us repeat what it means. Suppose A represents linear transformation T. Then w.r.t. eigenbasis $(\vec{v}_1, \dots, \vec{v}_n)$ T has diagonal matrix $B = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$. Said differently, one has the following formula representing the fact that matrix B is *similar* to matrix A:

$$A = \begin{bmatrix} | & & | \\ \vec{v}_1 & \dots & \vec{v}_n \\ | & & | \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \cdot \begin{bmatrix} | & & | \\ \vec{v}_1 & \dots & \vec{v}_n \\ | & & | \end{bmatrix}^{-1}$$

How to find eigenbasis, having eigenvalues

Suppose one found all eigenvalues $\lambda_1, \dots, \lambda_l$ of matrix $A_{n\times n}$, their algebraic multiplicities [1] are $almu(\lambda_i) = k_i$, and one has $k_1 + \dots + k_l = n$. The last condition is equivalent to $f_A(t) = \det(A - t \cdot I_n)$ being decomposed into n linear terms. This is the starting point for the second step in the algorithm above.

[1] *Algebraic multiplicity* of eigenvalue λ of matrix $A_{n\times n}$ is the maximal degree $k = almu(\lambda)$ such that $(\lambda - t)^k g(t) = f_A(t)$

One picks the first eigenvalue λ_1, and finds all corresponding eigenvectors. The way to do it is the following. Recall that "$\vec{v}$ is eigenvector corresponding to eigenvalue λ_1" means that we have

$$A\vec{v} = \lambda_1\vec{v} \iff (A - \lambda_1 I)\vec{v} = 0$$

Thus we can first consider the *eigenspace* of eigenvalue λ_1, which is a subspace $E_{\lambda_1} \subset \mathbb{R}^n$ defined by

$$E_{\lambda_1} = Ker(A - \lambda_1 I) = \{v \in \mathbb{R}^n | \ (A - \lambda_1 I)\vec{v} = 0\}$$

Said differently, eigenspace of λ_1 is the space of all corresponding eigenvectors. Because we are interested not in all of eigenvectors, but only in the eigenbasis, it is natural to find basis of E_{λ_1}, which we denote by $(\vec{v}_1, \ldots, \vec{v}_r)$. The way to do it is to use Gauss-Jordan elimination method for finding the basis of $Ker(A - \lambda_1 I) = E_{\lambda_1}$.

Because we got r vectors in the basis of E_{λ_1}, we have $dim(E_{\lambda_1}) = r$. This number has its special name:

Definition 1. *Geometric multiplicity* of eigenvalue λ of matrix $A_{n\times n}$ is by definition $dim(E_\lambda)$.

Proposition. *We always have $gemu(\lambda) \leqslant almu(\lambda)$.*

It turns out, that if $gemu(\lambda) < almu(\lambda)$, then diagonalization is not possible. Thus, returning to our process, having a basis $(\vec{v}_1, \ldots, \vec{v}_r)$ of $E_{\lambda_1} = Ker(A - \lambda_1 I)$, we hope that $r = almu(\lambda_1) = k_1$. If that is true we go to the next step: the same process for the eigenvalue λ_2.

This is what one does for every eigenvalue $\lambda_1, \ldots, \lambda_l$ hoping that on every step one has $\dim E_{\lambda_i} = k_i$. If that is the case, then in the end one will get exactly n linear independent eigenvectors, which will form an eigenbasis.

Example 1. We return to our example from the previous review session. There, for $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ we just guessed eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Let us find them by our algorithm. First, to find eigenvalues we find the roots of $f_A(t)$:

$$\det(A - tI_2) = \det \begin{bmatrix} -t & 1 \\ 1 & -t \end{bmatrix} = t^2 - 1 = (t-1)(t+1)$$

and thus ± 1 are the eigenvalues.

Now, let's find eigenspace E_{+1}:

$$Ker(A - 1I) = Ker \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} = span(\begin{bmatrix} 1 \\ 1 \end{bmatrix})$$

and we indeed see, that we can take $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as our first eigenvector, corresponding to eigenvalue 1.

Now, let's find eigenspace E_{-1}:

$$Ker(A - 1I) = Ker \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = span(\begin{bmatrix} 1 \\ -1 \end{bmatrix})$$

and we indeed see, that we can take $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ as our second eigenvector, corresponding to eigenvalue -1.

What can go wrong. As we said before, the problem in finding eigenvectors occurs when one has $gemu(\lambda) < almu(\lambda)$.

Example 2 (canonical example of not diagonalizable matrix)**.** Consider $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. One has $f_A(t) = (t-1)^2$, and so the only eigenvalue is $\lambda = 1$ with $almu(1) = 2$. Let's compute its eigenspace:

$$E_1 = Ker(A - 1I) = Ker \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = span(\begin{bmatrix} 1 \\ 0 \end{bmatrix})$$

is one dimensional! Thus $gemu(1) = 1 < 2 = almu(1)$, and one cannot find eigenbasis for this matrix.

Example 3. Is the following matrix diagonalizable?

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Remark. This problem in the process of diagonalization occurs both in the real and complex case.

This is the end of the description of algorithm for diagonalization.

Notice, that by definition of characteristic polynomial and Vieta's theorem one has

$$\det A = \lambda_1 \cdot \ldots \cdot \lambda_n$$

and

$$trace(A) = \lambda_1 + \ldots + \lambda_n$$

The following is also a useful theorem:

Theorem. *Suppose two square matrices A and B are similar (i.e. they represent the same linear transformation, just w.r.t different bases). Then the following are true:*

1. *Their characteristic polynomials are the same*

$$f_A(t) = f_B(t)$$

2. *They have the same eigenvalues, and also algebraic and geometric multiplicities of those.*

3. *They have the same ranks*

$$rk(A) = rk(B)$$

4. *They have the same determinant $\lambda_1 \cdot \ldots \cdot \lambda_n$ and trace $trace(A) = \lambda_1 + \ldots + \lambda_n$*

5. *Warning: eigenvectors can be different for A and B.*

7.5 Complex Eigenvalues

Here we cover the case when 2x2 matrix A has characteristic polynomial with no real roots. In this case we cannot diagonalize matrix A, but instead we can find a basis $(\vec{w}, \vec{v})$, w.r.t. which linear transformation is rotation+scaling, i.e. has the form $B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. Here is the recipe of how to find such basis.

Suppose $\lambda_{1,2} = a \pm ib$ are complex roots of $f_A(\lambda)$ (they are forced to be conjugate, because coefficients of $f_A(\lambda)$ are real). Then one finds a complex eigenvector $\vec{u} = \vec{w} + i\vec{v}$, corresponding to eigenvalue $a + ib$. This can be done by solving the system

$$(A - (a + bi)I_2)\vec{u} = 0$$

using Gauss-Jordan elimination method. Then the basis we were trying to find is just imaginary and real part of the eigenvector: $\vec{u} = \vec{v} + i\vec{w}$. The statement is that w.r.t. this basis linear transformation is given by rotation+scaling matrix

$$B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = S^{-1}AS$$

where change of basis matrix is $S = \begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix}$.

Example 4 (T/F question)**.** $A = \begin{bmatrix} 0 & 2 \\ -1 & 2 \end{bmatrix}$. Is it true that $\lim_{n\to\infty} A^n \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 0$?

Solution. False. The key is to understand what this matrix A does geometrically. For that one wants to find a basis, w.r.t. which matrix A is simpler to understand. Let's start by computing char polynomial:

$$f_A(t) = t^2 - 2t + 2$$

Its roots $\lambda_{1,2} = -1 \pm i = a + ib$ are complex. This means that even though A is not diagonalizable, with respect to some basis the linear transformation looks like

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}$$

which is rotation by $3\pi/4$ composed with scaling by $\sqrt{2}$. Thus A pushes every vector out of 0 by a factor of $\sqrt{2}$. From this geometric description it is clear that $\lim_{n\to\infty} A^n \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \infty \neq 0$.

We already solved the problem, but for educational purposes lets find a basis w.r.t. which our linear transformation is rotation+scaling:

$$E_{-1+i} = Ker \begin{bmatrix} -1-i & 2 \\ -1 & -1-i \end{bmatrix} = span(\begin{bmatrix} 1+i \\ -1 \end{bmatrix}) \implies \vec{u} = \begin{bmatrix} 1+i \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + i \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and so $\vec{w} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ form our desired basis. $\square$

True/False questions

Problem 1. All triangular matrices are diagonalizable.

Solution. False. Counterexample is $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ □

Problem 2. If $A = \begin{bmatrix} 7 & a & b \\ 0 & 7 & c \\ 0 & 0 & 7 \end{bmatrix}$ is diagonalizable, then a=b=c=0.

Solution. True. The char polynomial is $f_A(t) = (7-t)^3$. Thus one has only one eigenspace $E_7 = Ker(A - 7I_3) = \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}$. Because $almu(7) = 3$, for being able to diagonalize we need

$$gemu(7) = dim(Ker \begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}) = 3$$

This is only possible if $\begin{bmatrix} 0 & a & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{bmatrix}$ is a zero matrix, i.e. a=b=c=0. □

Problem 3. $f_A(t) = f_B(t) \implies$ A is similar to B.

Solution. False. Counterexample is $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. □

Problem 4. Any orthogonal 2x2 matrix is similar to rotation matrix.

Solution. False. Take any reflection matrix. It is orthogonal, and has eigenvalues ± 1, and thus is not similar to rotation matrix (which is either not diagonalizable, or is a rotation by π and so has one eigenvalue -1). □

Problem 5. A has no real eigenvalues $\implies$ $A^T A$ has no real eigenvalues.

Solution. False. Counterexample is a rotation by $\pi/2$, i.e. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Then $A^T A = I_2$ has real eigenvalue 1. □

Review and True/False questions, Bretscher 7.6,8.1

Review

7.6 Stability

Suppose we are given a discrete dynamical system in the usual form:

$$\begin{cases} \vec{x}(n+1) = A \cdot x(n) \\ \vec{x}(0) = \vec{x}_0 \end{cases}$$

The question we are going to explore is what happens with the trajectories as the time goes to ∞, i.e.

$$What\ is\ \lim_{t\to\infty} \vec{x}(t) = A^t \vec{x}_0?$$

The answer to that question in general depends on the matrix A, as well as on the initial value $\vec{x}_0$. But it turns out that in some cases we can say that the limit is common for all initial vectors.

Definition 1. Suppose we are given a discrete dynamical system $\vec{x}(n+1) = A \cdot x(n)$. Then $\vec{0}$ is called *stable equilibrium* of this dynamical system if $\lim_{t\to\infty} \vec{x}(t) = \vec{0}$ for every initial value $\vec{x}(0)$.

Let us consider 2x2 case. Remember, that we solved a dynamical system as long as the matrix is either [1. diagonalizable over reals] (end of section 7.2), or [2. has complex eigenvalues] (section 7.5) [1]. Let's recall what were the answers:

1. (accompanied with the illustration of phase portrait) A is diagonalizable over the reals. In this case, one first finds eigenvalues of A (possibly equal), and corresponding eigenvectors: λ_1, λ_2 and $\vec{v}_1, \vec{v}_2$. Then one decomposes an initial vector along the eigenbasis $\vec{x}(0) = c_1\vec{v}_1 + c_2\vec{v}_2$. Then the answer is

   $$\vec{x}(t) = c_1\lambda_1^t\vec{v}_1 + c_2\lambda_2^t\vec{v}_2$$

   Observation. Suppose $|\lambda_1| < 1$, $|\lambda_2| < 1$. Then we have a phase portrait where everything flows into the $\vec{0}$. More rigorously:

   $$\lim_{t\to\infty} \vec{x}(t) = \lim_{t\to\infty} c_1\lambda_1^t\vec{v}_1 + c_2\lambda_2^t\vec{v}_2 = \vec{0}$$

   for every initial value $\vec{x}(0)$. Thus $\vec{0}$ is stable equilibrium.

[1] Recall that these do not exhaust all the 2x2 cases. There is one more case, where the matrix is not diagonalizable over any numbers, for example the sheer matrix

2. (accompanied with the illustration of phase portrait) A has complex eigenvalues. In this case, one has that eigenvalues are conjugate $\lambda_1 = a + bi, \lambda_2 = a - bi$. One then finds the eigenvector $\vec{u} = \vec{v} + i\vec{w}$ for λ_1. Then it turns out that in the basis $\{\vec{w}, \vec{v}\}$ the matrix A becomes a rotation matrix

$$r \begin{bmatrix} cos(\alpha) & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = S^{-1}AS$$

(where $r = \sqrt{a^2 + b^2}$, $\alpha = arccos(\frac{a}{\sqrt{a^2+b^2}})$, and S is a change of basis matrix consisting of columns $\vec{w}, \vec{v}$)

Observation. Suppose $r < 1$. Then we have phase portrait (in the coordinates along $\vec{w}$ and $\vec{v}$) consisting of spirals, along which the trajectories go to $\vec{0}$. Thus:

$$\lim_{t \to \infty} \vec{x}(t) = \vec{0}$$

for every initial value $\vec{x}(0)$. Thus $\vec{0}$ is stable equilibrium.

Is there something common about those two cases? Yes! Notice, that in the second case one has

$$r = \sqrt{a^2 + b^2} = |a + bi| = |a - bi| = |\lambda 1| = |\lambda 2|$$

Thus both conditions in the two observations can be written as

$$|\lambda_1| < 1, |\lambda_2| < 1$$

without caring too much about those eigenvalues being real, or complex. In fact, one has much more general theorem:

Theorem. *Consider dynamical system $\vec{x}(n+1) = A \cdot x(n)$ (A is real matrix). Then $\vec{0}$ is stable equilibrium if and only if $|\lambda_i| < 1$ for all λ_i complex eigenvalues of A.*

8.1 Symmetric matrices

This is one of the most important results in linear algebra.

Theorem (Spectral theorem)**.**
Real matrix A has **orthonormal** *eigenbasis if and only if A is* **symmetric***, i.e. $A = A^T$. In short*

$$\text{"}symmetric \iff orthonormal\ eigenbasis\text{"}$$

Remark. From this one sees

$$\text{"}symmetric \implies diagonalizable\text{"}$$

because having an orthonormal basis is strictly stronger then having some eigenbasis.

Warning. The converse is false, i.e

$$\text{"}diagonalizable \not\Rightarrow symmetric\text{"}$$

Take for example matrix $\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$

A typical routine problem, which one faces in this chapter, is "how to find that orthonormal eigenbasis for symmetric matrices"? There two main steps.

The first step is to just find any eigenbasis, using the algorithm described in 7.1-7.3. Let us specialize for simplicity, and suppose that there were two eigenspaces: E_{λ_1} with basis $\vec{v}_1, \vec{v}_2$, and E_{λ_2} with basis $\vec{v}_3, \vec{v}_4, \vec{v}_5$.

In the second step we transform the vectors $\vec{v}_1, \ldots, \vec{v}_5$ s.t. they become orthonormal, but still are eigenvectors. The way to do it is to run Gram-Smidt process. But we cannot run G-S on all of the vectors, because then they will stop being eigenvectors. The way to do it is to run G-S separately for each eigenspace. First run G-S on the eigenspace E_{λ_1} for $\vec{v}_1, \vec{v}_2$ (get $\vec{u}_1, \vec{u}_2$), and then on E_{λ_2} for $\vec{v}_3, \vec{v}_4, \vec{v}_5$ (get $\vec{u}_3, \vec{u}_4, \vec{u}_5$). This way we ensure that 1) vectors in E_{λ_1} are orthonormal, and are still eigenvectors 2) vectors in E_{λ_3} are orthonormal, and are still eigenvectors. But why vectors from different eigenspaces should be orthonormal, for example $\vec{u}_1$ and $\vec{u}_4$? Turns out this comes for free, from the fact that the initial matrix was symmetric.

True/False questions

Let us first list the following list of useful facts:

Useful facts for T/F questions.

1. If A is orthogonal (i.e. $A^T = A^{-1}$), then all its eigenvalues (including complex ones) have absolute value one $|\lambda| = 1$. In particular, if the eigenvalue is real, it can only be $+1$ or -1. Complex eigenvalues lie on the unit circle.
2. Rotations by $\theta \neq k\pi$ on the plane $\mathbb{R}^2$ have no real eigenvalues, and their complex eigenvalues are $\cos(\theta) + i\sin(\theta)$
3. If $A_{n\times n}$ has real entries, and n is odd, then A has at least one real eigenvalue.
4. Characteristic polynomials are the same (and therefore eigenvalues too) for transpose matrices A and A^T.
5. Characteristic polynomials are the same (and therefore eigenvalues too) for similar matrices, i.e. for A and B s.t. $B = S^{-1}AS$ for invertible S.
6. Converse is false. For example the following two matrices have the same char. polynomial, but are not similar: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$
7. The property of being diagonalizable can break after the process of simplification to rref: $A \to rref(A)$. An example would be two matrices from the previous fact.

Problem 1. A is orthogonal matrix $\implies$ $A + A^{-1}$ is diagonalizable.

Solution. True. A is orthogonal, so $A^T = A^{-1}$. Thus $A + A^{-1} = A + A^T$ which is a symmetric matrix for any A. $\square$

Problem 2. A is symmetric, and $A^2 = 0$. Then $A = 0$.

Solution. True. Because A is symmetric, it is diagonalizable. So it definitely has real eigenvalues. The relevant question one should ask now is "what are the possible options for eigenvalues of A?".

Remember from previous problems, of λ is an eigenvalue of A, then λ^2 is an eigenvalue of A^2. But $A^2 = 0$, and the only eigenvalue 0-matrix has is 0. Thus we get that $\lambda = 0 \implies \lambda = 0$ for any eigenvalue of A. Thus our matrix A is similar to a diagonal matrix with eigenvalues (which are 0) on the diagonal, i.e. to a 0-matrix. This means that A is itself a 0-matrix, q.e.d. $\square$

Problem 3. A is symmetric $\implies A^3\vec{x} = A^2\vec{b}$ has a solution always.

Solution. True. Remember the section about the least square solution. If one has a system of equations $A\vec{x} = \vec{b}$, it might happen that it does not have solutions. But $A^T A\vec{x} = A\vec{b}$ always has solutions. Now we use that A is symmetric:

$$A^T A\vec{x} = A\vec{b} \iff A^2\vec{x} = A\vec{b} \implies A^3\vec{x} = A^2\vec{b}$$

Because the equation on the left has a solution, the equation on the right also must have this solution (it is a consequence of the equation on the left). Q.e.d. $\square$

Problem 4. A is symmetric 4x4, $f_A(\lambda) = \lambda^2(\lambda - 1)(\lambda - 2)$. Then $Ker(A)$ is 2-dim.

Solution. True. A is symmetric, so it is diagonalizable, and so $almu = gemu$ always. Thus we have $2 = almu(0) = gemu(0) =$ dimension of eigenspace corresponding to 0 eigenvalue $= Ker(A - 0 \cdot I_2) = Ker(A)$. $\square$

Problem 5. $A = \begin{bmatrix} -1 & 3 & -1 \\ 0 & 1 & 0 \\ 2 & -3 & 2 \end{bmatrix}$ is orthogonal projection.

Solution. False. Orthogonal projection matrices are always symmetric. $\square$

Problem 6. A is orthogonal, and diagonalizable over $\mathbb{R}$. Then A is either $\pm I_2$, or reflection.

Solution. True. Understand what A can do to the eigenbasis (the key restriction is that A preserves lengths, and thus can only either preserve, or flip the eigenbasis). $\square$

Problem 7. A is 3x3 orthogonal matrix with det(A)=1. Then 1 is an eigenvalue of A.

Solution. True. det(A)=1 implies $\lambda_1\lambda_2\lambda_3 = 1$. A is orthogonal implies $|\lambda_i| = 1$. From this the statement follows. $\square$

Problem 8. There exists 3x3 matrix A with det=3 and entries $=\pm 1$.

Solution. False. If all entries are odd, then determinant will be even, so cannot be 3. $\square$

Review and True/False questions, Bretscher 8.2,8.3

Review

8.2 Quadratic forms

Recall that the main point from 8.1 was the spectral theorem: "real matrix A is symmetric if and only if it has orthonormal eigenbasis, i.e. it is diagonalizable w.r.t. orthonormal basis". With this powerful tool we will be able to understand quadratic forms.

Definition 1. Quadratic form is a function $q : \mathbb{R}^n \to \mathbb{R}$ which can be written in the form

$$q(\vec{x}) = \sum_{i \leqslant j} a_{ij} x_i x_j = \vec{x}^T A \vec{x}$$

Example 1. $q(x_1, x_2) = 2x_1^2 + 6x_1x_2 + 5x_2^2 = [x_1, x_2] \cdot \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Remark. Notice, that in the quadratic form we had coefficient 6, but in the matrix we have $3 = 6/2$. Also, notice that the matrix A coming from a quadratic form is always symmetric.

Goal. Understand quadratic forms via change of basis. In particular we would like to understand the shape of the level sets, for example $q(\vec{x}) = 1$.

Method. Diagonalization. The point is that matrix A coming from quad. form is always symmetric, and therefore diagonalizable. Here is the step by step procedure:

1. We are given a quadratic form $q(\vec{x}) = \vec{x}^T A \vec{x}$.

2. A is symmetric, so we can diagonalize it, and moreover make its eigenbasis to be orthonormal. I.e.

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

   where $\lambda_1, \ldots, \lambda_n$ are eigenvalues of A (some may be equal), and columns of S are vectors from orthonormal eigenbasis $\{\vec{u}_1, \ldots, \vec{u}_n\}$. In particular, $A\vec{u}_i = \lambda_i \vec{u}_i$.

   Remark. Because $\{\vec{u}_1, \ldots, \vec{u}_n\}$ are orthonormal, it is particularly easy to find S^{-1}, namely $S^{-1} = S^T$. This fact is responsible for the following step.

3. Suppose $\vec{x} = c_1\vec{u}_1 + \cdots + c_n\vec{u}_n$. I.e. $c_1, \ldots, c_n$ are coordinates of $\vec{x}$ in the basis $\{\vec{u}_1, \ldots, \vec{u}_n\}$ [1]. Then
$$q(\vec{x}) = \lambda_1 c_1^2 + \cdots \lambda_n c_n^2$$

4. Principal axes of quadratic form $q(\vec{x}) = \vec{x}^T A\vec{x}$ are the lines along the vectors $\vec{u}_i$ (which form orthonormal eigenbasis of A).

   If $\lambda_i > 0$, then $\{q(\vec{x}) = 1\}$ intersects $\vec{u}_i$ axis in the 2 points of abs value $\frac{1}{\sqrt{\lambda_i}}$.

5. Shape: (accompanied with illustrations)

   (a) Case $\lambda_i > 0$ for all i. Then $\{q(\vec{x}) = 1\}$ is ellipsoid, intersecting principal axis in the points specified above.

   (b) Case $\lambda_i < 0$ for all i. Then $\{q(\vec{x}) = 1\}$ is empty.

   (c) Case λ_i have different signs. Then $\{q(\vec{x}) = 1\}$ is unbounded (in the 2x2 case it is hyperbola).

Definition 2. Quadratic form $q(\vec{x})$ is called positive definite if $q(\vec{x}) > 0$ for all $\vec{x} \neq 0$. Equivalently if all $\lambda_i > 0$. Equivalently if $\{q(\vec{x}) = 1\}$ is an ellipsoid.

How to quickly check if the quadratic form is positive definite? (without solving $f_A(\lambda) = 0$)

Theorem. *Quadratic form $q(\vec{x}) = \vec{x}^T A\vec{x}$ is positive definite if and only if determinants of all principal submatrices A_i are < 0:*

$$\begin{array}{c} A_1 \\ A_2 \\ A_3 \end{array} \begin{pmatrix} 8 & 1 & 6 \\ 8 & 5 & 7 \\ 8 & 9 & 5 \end{pmatrix}$$

8.3 Singular values and SVD

Suppose linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ is represented by $A_{n\times m}$ (w.r.t. the standard basis of course). Notice that in this chapter we are studying non-square matrices too.

Goals. **I**. Describe the image of a unit sphere $S = \{x_1^2 + \cdots + x_m^2\} \subset \mathbb{R}^m$ under the linear transformation, i.e. what is $T(S) \subset R^n$?

II. Decompose matrix A in the following way:

$$A_{n\times m} = U_{n\times n} \cdot \Sigma_{n\times m} \cdot V^T_{m\times m}$$

where U and V are orthogonal (i.e. columns are orthonormal) square matrices, and Σ is a matrix of the same size as A, having everywhere 0 coefficients except its 45° diagonal, and on that diagonal having coefficients $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_m \geqslant 0$, which are called singular values. [2]

[1] one can find those coordinates by solving a linear system of equations for $\vec{x} = c_1\vec{u}_1 + \cdots + c_n\vec{u}_n$ for c_i).

[2] See Figure 4 on page 409 of the book for geometric interpretation of SVD.

Algorithm. **1**. Compute eigenvalues of $(A^TA)_{m\times m} = B_{m\times m}$. Note that this is possible, as A^TA is always symmetric. The eigenvalues will be all non-negative:

$$\text{eigenvalues}: \ \sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_m \geqslant 0$$

From eigenvalues by taking square root we get the <u>singular values</u> of matrix $A_{n\times m}$:

$$\text{singular values: } \sigma_1 = \sqrt{\lambda_1} \geqslant \sigma_2 = \sqrt{\lambda_2} \geqslant \ldots \geqslant \sigma_m = \sqrt{\lambda_m} \geqslant 0$$

Note, that there will be exactly $rk(A) = k$ many non-zero eigenvalues (and therefore singular values), all others will be 0.

2. Find orthonormal eigenbasis of $(A^TA)_{m\times m} = B_{m\times m}$ (possible because it is symmetric)

$$\{\vec{v}_1, \ldots, \vec{v}_m\}$$

Recall that the way to do it is to first find some eigenbasis for A^TA, and then for each eigenspace run Gram-Schmidt process process for its eigenvectors.

3. Answer to the Goal I. Recall that we interested in the image of the unit circle under $A_{n\times m}$.

- Draw axes along $A\vec{v}_1, \ldots, A\vec{v}_m$. We will draw an ellipsoid along those axes. The lengths of the axes of that ellipsoid will be $\sigma_1, \ldots, \sigma_m$. The points of intersections of axes with ellipsoid will be non-zero $\pm A\vec{v}_i$.
- If there are no 0's among those singular values, then the answer is that image of the unit sphere is an ellipsoid with those axes $A\vec{v}_1, \ldots, A\vec{v}_m$ and lengths of axes $\sigma_1, \ldots, \sigma_m$.

  If there are zeroes among those vectors (i.e. $rk(A) < m$), it means that our sphere was collapsed at some point. Then the answer is the following: ignore those 0 vectors and sing values, draw an ellipsoid along the non-zero vectors, and then the answer is the <u>full interior of that ellipse</u>.

  (accompanied with illustration) For example it is possible that our matrix was $A_{3\times 3}$, giving a map from $\mathbb{R}^3$ to $\mathbb{R}^3$. Suppose it has rank=2. Then the image of the unit sphere will not be an ellipsoid, but rather it will be a flat interior of an ellipse (because ellipsoid was collapsed).

4. Answer to the Goal II. Recall that we want

$$A_{n\times m} = U_{n\times n} \cdot \Sigma_{n\times m} \cdot V^T_{m\times m}$$

- $\Sigma_{n\times m}$ is easy to find. Just take $n \times m$ matrix Σ, and put zeros everywhere except 45° diagonal, where put the singular values $\sigma_1, \ldots, \sigma_m$.
- $V = \begin{bmatrix} | & & | \\ \vec{v}_1 & \ldots & \vec{v}_m \\ | & & | \end{bmatrix}$, i.e. $V^T = \begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_m & - \end{bmatrix}$.

- It is left to find $U_{n\times n}$. First $k = rk(A) =$ (number of non-zero σ_i) vectors are

$$\frac{1}{\sigma_1}A\vec{v}_1, \ldots, \frac{1}{\sigma_k}A\vec{v}_k$$

  To find the rest just spot additional vectors $\vec{u}_{k+1}, \ldots, \vec{u}_n$ s.t. $\{\vec{u}_1, \ldots, \vec{u}_k, \vec{u}_{k+1}, \ldots, \vec{u}_n\}$ is orthonormal basis of $\mathbb{R}^n$.

  If it is not easy to guess, here are two recipes for how to find the rest $\vec{u}_{k+1}, \ldots, \vec{u}_n$ vectors:

  - (recommended way) find orthonormal basis of $Ker(A^T)$ (first find any basis, then run Gram-Schmidt process).
  - It turns out that one can find $\vec{u}_{k+1}, \ldots, \vec{u}_n$ by computing eigenvectors corresponding to eigenvalue 0 of matrix $C = AA^T$.

  Remark. Vectors $\vec{u}_{k+1}, \ldots, \vec{u}_n$ are not unique. Question: what about vectors $\vec{u}_1, \ldots, \vec{u}_k$, i.e. vectors $\vec{v}_1, \ldots, \vec{v}_k$?

True/False questions

I forgot to say a very useful fact last time: orthogonal projections always have symmetric matrices. Question: why? Because they have orthonormal eigenbasis.

Problem 1. $A = \begin{bmatrix} -1 & 3 & -1 \\ 0 & 1 & 0 \\ 2 & -3 & 2 \end{bmatrix}$ is orthogonal projection.

Solution. False. Orthogonal projection matrices are always symmetric. □

Problem 2. The product of two quadratic forms is also a quadratic form.

Solution. False. Take $q_1(x_1, x_2) = x_1^2$ and $q_2(x_1, x_2) = x_1^2$. Then $q_1 q_2 = x_1^4$ is not a quadratic form. □

Problem 3. The sum of quadratic forms is quadratic.

Solution. True. □

Problem 4. $x^2 + 6xy + 7y^2 = 2017$ is an ellipse in $\mathbb{R}^2$.

Solution. False. Compute the determinants of principal submatrices, and they have different signs. Thus by the theorem we have that $x^2+6xy+7y^2 = 1$ is not an ellipse (it is unbounded). Thus $x^2 + 6xy + 7y^2 = 2017$, which differ from $x^2 + 6xy + 7y^2 = 1$ by scaling (by what factor?), is not an ellipse too. □

The following problem is about the question "in how many ways one can diagonalize orthonormally?".

Problem 5. $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$. Then there exists exactly four orthogonal 2x2 matrices S s.t. $S^{-1}AS$ is diagonal.

Solution. If one denotes columns of S by $\vec{v}_1, \vec{v}_2$, then the task is to compute the number of orthonormal bases $\vec{v}_1, \vec{v}_2$, w.r.t. which A is diagonal.

First of all, note that eigenvalues $\lambda_{1,2}$ are different, otherwise it would be a scaling matrix, which is not the case. Another approach is to just compute $\lambda_{1,2} = 2 \pm \sqrt{5}$.

This means that there are two 1-dim. eigenspaces E_{λ_1} and E_{λ_2}. We have $\vec{v}_1 \in E_{\lambda_1}$ and its length should be 1 (so that the basis is orthonormal). So there are two such vectors, denote $\pm\vec{u}_1$. We have $\vec{v}_2 \in E_{\lambda_2}$ and its length should be 1, denote two choices by $\pm\vec{u}_2$.

Thus it looks like we have exactly 4 choices for: $S = [\pm\vec{u}_1, \pm\vec{u}_2]$. But in fact we have 4 more choices, because we change the order of eigenvalues, so the last 4 choices are $S = [\pm\vec{u}_2, \pm\vec{u}_1]$.

So we have 8 choices, not 4. Thus the answer is False. □

Here are problems which illustrate, that singular values of matrix A (which square roots of eigenvalues of $A^T A$) behave very differently from eigenvalues of A.

Problem 6. Similar matrices have the same singular values.

Solution. False. Take $A = \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix}$ and $A = \begin{bmatrix} 0 & 4 \\ 0 & 3 \end{bmatrix}$. □

Problem 7. Singular values of $n \times n$ matrix are square roots o eigenvalues.

Solution. False. Take matrix $A = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$. Then it has $\lambda_{1,2} = \pm\sqrt{2}$. Whereas $A^T A = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$, and so singular values are 2 and 1. □

Problem 8. A is 2x2 matrix with singular values $\sigma_1, \sigma_2 = 1, 2$. Then A^2 has singular values 1 and 4.

Solution. False. Take the same matrix $A = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$. We have $A^T A = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$, and so singular values are 2 and 1. But $A^2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ has eigenvalues 2 and 2. □

Review and True/False questions, Bretscher 9.1,9.2

Review

9.1 Continuous dynamical systems

First, how do we differentiate vectors, matrices? Entry-wise. For example

$$\frac{d}{dt}\left(\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}\right) = \begin{bmatrix} \frac{d}{dt}(x_1(t)) \\ \frac{d}{dt}(x_2(t)) \end{bmatrix}$$

Continuous dynamical system is described via the following system

$$\begin{cases} \frac{d}{dt}(\vec{x}(t)) = A \cdot x(t) \\ \vec{x}(0) = \vec{x}_0 \end{cases}$$

At each point we know velocity of the trajectory, as opposed to how to jump to the next vector in the discrete dyn. system case.

Example 1. $\frac{d}{dt}(\vec{x}(t)) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot x(t)$

Here it is easy to spot the trajectories $\vec{x}(t) = r(\cos(t), \sin(t))$, and draw a phase portrait: (illustration on the blackboard).

The question is how to solve continuous dynamical systems in general, i.e. how to find a trajectory depending on A and initial value. In this section we start with the case when A is diagonalizable over $\mathbb{R}$.

Theorem. *Suppose A is 2x2, diagonalizable over $\mathbb{R}$. Suppose λ_1, λ_2 are its eigenvalues (possibly equal, we count with algebraic multiplicities), $\vec{v}_1, \vec{v}_2$ are its eigenvectors. Consider a continuous dynamical system*

$$\begin{cases} \frac{d}{dt}(\vec{x}(t)) = A \cdot x(t) \\ \vec{x}(0) = \vec{x}_0 \end{cases}$$

Then the following gives a method to compute trajectories:

1. *Case $\vec{x}_0 = \vec{v}_1$. Then we reduce the problem to 1-dim case $\dot{x}(t) = \lambda_1 x(t), x(0) = 1$ along the line along $\vec{v}_1$, and we have a solution*

$$\vec{x}(t) = e^{\lambda_1 t}\vec{v}_1$$

2. *Case* $\vec{x}_0 = \vec{v}_2$. *Analogously we have*

$$\vec{x}(t) = e^{\lambda_2 t}\vec{v}_2$$

3. *General case* $\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2$. *Then by linearity one gets*

$$\vec{x}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_2 t}\vec{v}_2$$

Thus here are the steps for solving continuous dyn. system in the diagonalizable case: 1) find eigenvalues 2) find eigenbasis 3) decompose the initial vector along the eigenbasis (via solving the linear system $\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2$ for c_1 and c_2) 4) write down the formula $\vec{x}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_2 t}\vec{v}_2$.

Remark. The formula for $n \times n$ case is completely analogous.

The way to draw a phase portray is to draw eigenspaces, understand what happens along them (depends on the sign of λ_i), and then fill in the plane according to those flows.

9.2 Case 2x2 with real entries, complex eigenvalues

In this case we have to work a bit more. We want to solve

$$\begin{cases} \frac{d}{dt}(\vec{x}(t)) = A \cdot x(t) \\ \vec{x}(0) = \vec{x}_0 \end{cases}$$

where $A_{2\times 2}$ has complex eigenvalues.

One can just do what we did in the previous section, but over complex numbers. Namely:

- Find complex eigenvalues λ_1,λ_2 of the matrix A.
- Find the corresponding complex eigenvectors $\vec{v}_1$,$\vec{v}_2$.
- Decompose the initial vector along the eigenbasis, solving the linear system $\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2$ over complex numbers, getting a complex solution c_1, c_2
- Write the formula
$$\vec{x}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_2 t}\vec{v}_2$$

Because λ_1 is a complex number, it is not quite clear what $e^{\lambda_1 t}$ means. Turns out that one can understand what does it mean to raise to the complex power, and the main things to remember is the Euler's formula:

$$e^{p+iq} = e^p(\cos(q) + i\sin(q))$$

Now, a fair question to ask is "why our solutions in the end will be real?". Because, as one notices from the Euler's formula, there will be imaginary parts showing up in the answer. Turns out that magically, if one started with matrix A with real coefficients, and real vector $\vec{x}_0$ as initial value, then all the imaginary part will cancel and one eventually will get a real solution.

Example 2. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $\vec{x}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We guessed already that the solutions are $\vec{x}(t) = r(\cos(t), \sin(t))$, but let's try to use the method above.

Eigenvalues: $\lambda_{1,2} = \pm i$. Corresponding eigenvectors: $\vec{v}_{1,2} = \begin{bmatrix} 1 \\ \mp i \end{bmatrix}$. The decomposition of initial value is $\vec{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1/2\vec{v}_1 + 1/2\vec{v}_2$. Let's write the formula now:

$$\vec{x}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + c_2 e^{\lambda_2 t}\vec{v}_2 = 1/2e^{it}\begin{bmatrix} 1 \\ -i \end{bmatrix} + 1/2e^{-it}\begin{bmatrix} 1 \\ i \end{bmatrix} =$$

$$= (1/2)\Big((\cos(t) + i\sin(t))\begin{bmatrix} 1 \\ -i \end{bmatrix} + (\cos(t) - i\sin(t))\begin{bmatrix} 1 \\ i \end{bmatrix}\Big) =$$

$$= (1/2)\begin{bmatrix} \cos(t) + i\sin(t) + \cos(t) - i\sin(t) \\ -i\cos(t) + \sin(t) + i\cos(t) + \sin(t) \end{bmatrix} = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$

In order to avoid this cancellation of imaginary terms, there is a following theorem, which provides a shortcut to the answer:

Theorem. *Suppose $A_{2\times 2}$ is a matrix with real entries, but complex eigenvalues. We are interested in solving*

$$\begin{cases} \frac{d}{dt}(\vec{x}(t)) = A \cdot x(t) \\ \vec{x}(0) = \vec{x}_0 \end{cases}$$

Denote $\lambda_1 = p + iq$, $\lambda_2 = p - iq$, and $\vec{u} = \vec{v} + i\vec{w}$ is the complex eigenvector for λ_1. Consider also a matrix $S = [\vec{w}, \vec{v}]$. Then the solution is

$$\vec{x}(t) = e^{pt} \cdot S \cdot \begin{bmatrix} \cos(qt) & \sin(qt) \\ \sin(qt) & \cos(qt) \end{bmatrix} \cdot S^{-1}x_0$$

Trajectories are

$$\begin{cases} (p = 0) & \Longrightarrow & \textit{ellipses} \\ (p > 0) & \Longrightarrow & \textit{outward spirals} \\ (p < 0) & \Longrightarrow & \textit{inward spirals} \end{cases}$$

The angle period (in case of an ellipse this is an orbit period) is $\frac{2\pi}{q}$.

Remark. Note, that basis $\vec{w}, \vec{v}$ is precisely the basis in which the matrix A looks like a rotation+scaling matrix. But nevertheless, the formula for the solution in continuous case is very different from the formula in the discrete dyn system case.

The following table clarifies the difference between the formulas for solutions for discrete and continuous dynamical system.

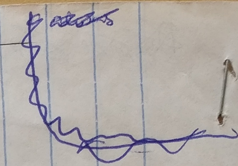
Dyn. ~~systems~~ systems

for phase portraits see page 438 of book.

	dyscrete dynamical systems	continous dynamical systems
diagonalizable over real	$\begin{cases} \overline{x}(t+1) = A\overline{x}(t) \\ x(0) = x_0 \end{cases}$, $\lambda_1 \dots \lambda_n$, $v_1 \dots v_n$ — need to find them; $x_0 = c_1\overline{v}_1 + \dots + c_n\overline{v}_n$ $\overline{x}(t) = A^t\overline{x}_0 = c_1\lambda_1^t\overline{v}_1 + \dots + c_n\lambda_n^t\overline{v}_n$ (by the way $S^{-1}AS = \begin{bmatrix}\lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}$, $S = [\overline{v}_1 \dots \overline{v}_n]$)	$\begin{cases} \frac{d\overline{x}}{dt} = Ax(t) \\ \overline{x}(0) = \overline{x}_0 \end{cases}$, $\lambda_1 \dots \lambda_n$, $v_1 \dots v_n$; $x_0 = c_1\overline{v}_1 + \dots + c_n\overline{v}_n$ $\overline{x}(t) = c_1 e^{\lambda_1 t}\overline{v}_1 + \dots + c_n e^{\lambda_n t}\overline{v}_n$
diagonalizable over $\mathbb{C}$ solution is real if A has real entries	$\begin{cases} \overline{x}(t+1) = Ax(t) \\ \overline{x}(0) = \overline{x}_0 \end{cases}$ ← complex; $\lambda_1 \dots \lambda_n$ — complex eigenvalues; $v_1 \dots v_n$ — complex eigenvectors; $\overline{x}_0 = c_1\overline{v}_1 + \dots + c_n\overline{v}_n$ ← complex → $x(t) = A^t\overline{x}_0 = c_1\lambda_1^t\overline{v}_1 + \dots + c_n\lambda_n^t\overline{v}_n$	$\begin{cases} \frac{d\overline{x}}{dt} = Ax(t) \\ \overline{x}(0) = x_0 \end{cases}$ complex; $\lambda_1 \dots \lambda_n$, $\overline{v}_1 \dots \overline{v}_n$ — complex; $x_0 = c_1\overline{v}_1 + \dots + c_n\overline{v}_n$ $x(t) = c_1 e^{\lambda_1 t}\overline{v}_1 + \dots + c_n e^{\lambda_n t}\overline{v}_n$
stability in those two cases	$\overline{0}$ is stable iff ~~modulus of eigenvalues are less than 1~~ $\lvert\lambda_i\rvert < 1$ for all λ_i (in both real and complex case)	$\overline{0}$ is stable iff $\lambda_i < 0$ for all λ_i (real case); $Re\,\lambda_i < 0$ for all λ_i (complex case)
2x2 case, when A is similar to rotation (will find real solutions) complex case when A has real entries, But complex λ_i's	$\begin{cases} \overline{x}(t+1) = A\overline{x}(t) \\ \overline{x}(0) = x_0 \end{cases}$; $\lambda_{1,2} = p \pm iq = r(\cos\alpha \pm i\sin\alpha)$, $r = \sqrt{p^2+q^2}$ $\overline{v} + i\overline{w}$ — eigenvector for λ_1 then $A = S\begin{bmatrix}\cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha\end{bmatrix}S^{-1}$, $S = [\overline{w}\ \overline{v}]$, and so $= S\begin{bmatrix}p & -q \\ q & p\end{bmatrix}S^{-1}$ $\overline{x}(t) = r^t S\begin{bmatrix}\cos t\alpha & -\sin t\alpha \\ \sin t\alpha & \cos t\alpha\end{bmatrix}S^{-1}\vec{x}_0$ $\boxed{\overline{0} \text{ is stable} \iff r < 1}$ (in this case)	$\begin{cases} \frac{d\overline{x}}{dt} = A\overline{x}(t) \\ \overline{x}(0) = \overline{x}_0 \end{cases}$; $\lambda_{1,2} = p \pm iq$; $\overline{v} + i\overline{w}$ — eigenvector for $p + iq$ $A = S\begin{bmatrix}p & -q \\ q & p\end{bmatrix}S^{-1}$, $S = [\overline{w}\ \overline{v}]$, and so $x(t) = e^{pt} S\begin{bmatrix}\cos(qt) & -\sin(qt) \\ \sin(qt) & \cos(qt)\end{bmatrix}\underbrace{S^{-1}\vec{x}_0}_{=\begin{bmatrix}c_1 \\ c_2\end{bmatrix}}$, where $\overline{x}_0 = c_1\overline{w} + c_2\overline{v}$ notice the difference! $\overline{0}$ is stable iff $tr\,A < 0$, $\det A > 0$ (in all ~~continuous~~ real 2x2 cases)

trajectories: ellipses if $p = 0$
- spirals outward $p > 0$
- spiral inward $p < 0$

period of trajectory: $\frac{2\pi}{q}$



True/False questions

Problem 1. Suppose A is 2x2 symmetric matrix with distinct eigenvalues $\alpha > \beta > 1$. Suppose $\vec{x}(t) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is a trajectory of a continuous dynamical system

$$\begin{cases} \frac{d}{dt}(\vec{x}(t)) = A \cdot x(t) \\ \vec{x}(0) = \begin{bmatrix} 0 \\ 4.9 \end{bmatrix} \end{cases}$$

For each of the following statements determine if it is True, False, or Undetermined (by the given information):

a) $lim_{t\to\infty}|\vec{x}(t)| = +\infty$

b) $lim_{t\to\infty}x_1(t) = +\infty$

c) $lim_{t\to\infty}x_1(t) = 0$

d) $lim_{t\to\infty}x_1(t) = -\infty$

e) $lim_{t\to\infty}x_2(t) = +\infty$

f) $lim_{t\to\infty}x_2(t) = 4.9$

g) $\vec{x}(t)$ is asymptotic to the line $x_1 = x_2$ as $t \to \infty$.

Solution. First of all, A is symmetric and has distinct eigenvalues, so it is definitely diagonalizable, let's say with respect to the eigenbasis $\vec{v}_1, \vec{v}_2$. Moreover $\alpha > \beta > 1 > 0$, and so the phase portrait looks like (illustration on the blackboard) a bunch of trajectories coming out of $\vec{0}$, where the straight trajectories are along $\vec{v}_1$ and $\vec{v}_2$, and the others a tangent to one of the coordinates (question: to what coordinate are they tangent?). In order to answer the questions, one needs to understand what are possible positions of vector $\vec{x}(0) = \begin{bmatrix} 0 \\ 4.9 \end{bmatrix}$ w.r.t. this phase portrait. All the following answers can be proved rigorously by the formula for the solution. Here we will use a geometric argument using the phase portrait.

a) T. This is simple, all the trajectories go to ∞ except the zero trajectory, but $\vec{x}(0) \neq \vec{0}$, so it will go to ∞, and so the abs value will need to go to ∞.

b) U. This is trickier. The point is that even if the vector goes to ∞ on the plane, one of its coordinates can actually stay the same (for example of $\vec{v}_1 = \vec{x}(0)$. Or it can go to $\pm\infty$. (see this by placing $\vec{x}(0)$ in the appropriate position on the phase portrait) Thus all the answers to this and the next two questions are "Undetermined".

c) U.

d) U.

e) T. Here one actually can see that no matter where you put $\vec{x}(0)$ w.r.t. to the phase portrait, it will go to $+\infty$

f) F. Follows from previous question.

g) F. It can be seen that the slope of the trajectories stabilizes to one of the eigenspaces, but there are no asymptotics, because the coordinate w.r.t. eigenspaces will both go to $\pm\infty$. Unless the trajectory is the eigenspace itself, but in this case we would have to have $0 = 4.9$, which is not true.

$\square$

Problem 2. If A is diagonalizable, then e^A is invertible.

Solution. First of all, how to define e^A? For diagonalizable case it is done via the following method: $D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = S^{-1}AS \implies A = SDS^{-1}$ and we define

$$e^A = Se^DS^{-1} = S \begin{bmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{bmatrix} S^{-1}$$

All matrices here are invertible (cause $e^{\lambda_i} \neq 0$), so the product is also invertible, q.e.d. $\square$

Problem 3. A is symmetric, $x(t)$ is a solution to $\begin{cases} \frac{d}{dt}(\vec{x}(t)) = A \cdot x(t) \\ \vec{x}(0) = \vec{x}_0 \end{cases}$. Suppose there exists initial vector $\vec{x}_0 \neq 0$ s.t. $\vec{x}(50) = \vec{x}_0$. Then it can happen that A is invertible.

Solution. False. Because A is symmetric, we know it is diagonalizable, and so the formula is

$$\vec{x}(50) = c_1e^{\lambda_1 50}\vec{v}_1 + \ldots + c_ne^{\lambda_n 50}\vec{v}_n$$

where $\vec{x}_0 = c_1\vec{v}_1 + \ldots + c_n\vec{v}_n$. For $\vec{x}(50) = \vec{x}_0$ we should have

$$c_ie^{\lambda_i 50} = c_i,$$

and so either all $c_i = 0$, which is impossible because $\vec{x}_0 \neq 0$, or there is some $c_k \neq 0$ and then $\lambda_k = 0$, which proves that A is non-invertible. $\square$

Problem 4. Take $A = \begin{bmatrix} -3 & -5 \\ -4 & 180 \end{bmatrix}$. Then for a continuous dynamical system $\frac{d}{dt}(\vec{x}(t)) = A \cdot x(t)$ we have that $\vec{0}$ is stable equilibrium.

Solution. False. Because we are in 2x2 continuous case, we can apply the criterion $tr(A) < 0, det(a) > 0 \iff \vec{0}$ is stable. But the criterion fails in our case, the trace is positive. $\square$

Problem 5. 2x2 matrix A has two distinct negative eigenvalues. Then $\vec{0}$ is stable for discrete dynamical system $\vec{x}(t+1) = A\vec{x}(t)$.

Solution. False. Because 2x2 matrix A has two distinct eigenvalues, it is diagonalizable. The criterion in the discrete diagonalizable case is that $|\lambda_i| < 1$, not $\lambda_i < 0$ (which is the criterion for continuous case). Thus, for example case where $\lambda_1 = -20$ would be a counterexample to the statement. $\square$