Hour to hour syllabus for Math21b, Spring 2018

Abstract

This is an outline of the lectures for the Spring 2018 semester. The section numbers refer to the book of Otto Bretscher, "Linear algebra with applications". The text is not mandatory. Homework is distributed on handouts. There might be tiny adaptations to the syllabus.

Week 0, Jan 25-Jan 26

1. Lecture: Introduction to linear systems, Section 1.1

Before introducing **Gauss-Jordan elimination**, we look informally at some **systems of linear equations**. How do such systems arise from applications and how can one solve them with 'ad hoc' methods like combining equations or by successive eliminating variables using substitution until only one variable is left? This first lecture allows us to introduce ourselves and the course assistant as well as let students fill out the preference sheet (which the course assistant will use for selecting a problem session time). Note that this is still shopping period.

Week 1, Jan 29-Feb 2

2. Lecture: Gauss-Jordan elimination, Section 1.2

Having rewritten systems of linear equations using **matrices** we introduce **Gauss-Jordan elimination steps**: scaling of rows, swapping rows or subtracting a multiple of a row to an other row. We also want to discuss an example, where one has more than one solution or where one has no solution. Unlike in multi-variable calculus, we distinguish between **column vectors** and **row vectors**. As most objects are vectors, we leave away the arrow when writing vectors. Column vectors are $n \times 1$ matrices, and row vectors are $1 \times m$ matrices. A general $n \times m$ matrix has m columns and n rows. The output of the Gauss-Jordan elimination is a matrix $\mathrm{rref}(A)$ which is in row reduced echelon form: the first nonzero entry in each row is 1, called **leading** 1, every column with a leading 1 has no other nonzero elements and every row above a row with a leading 1 has a leading 1 to the left.

3. Lecture: On solutions of linear systems, Section 1.3

How many solutions does a system of linear equations have? We see in this lecture that there are three possibilities: there can be exactly one solution, there can be no solution or there are infinitely many solutions. In small dimensions, this can be visualized and explained geometrically. We also learn to determine in which of the three case we are in, using Gauss-Jordan elimination or by looking at the **rank** of the matrix A and the **augmented matrix** $B = [A|b]$. We also mention that one can see a system of linear equations $Ax = b$ in two different ways: the **column picture** tells that $b = x_1 v_1 + \cdots + x_n v_n$ is a sum of column vectors v_i of the matrix A, the **row picture** tells that the dot product of the row vectors w_j with x are the components $w_j \cdot x = b_j$ of b.

4. Lecture: Linear transformation, Section 2.1

Geometric and algebraic descriptions of linear transformations are now linked. Linear transformations are introduced formally as transformations $T(x) = Ax$, where A is a matrix. We learn how to distinguish between linear and nonlinear, linear and affine transformations. The transformation $T(x) = x + 5$ for example is not linear because 0 is not mapped to 0. We characterize linear transformations on R^n by three properties: $T(0) = 0$, $T(x + y) = T(x) + T(y)$ and $T(sx) = sT(x)$. This assures compatibility with the additive structure on R^n. We see many examples of linear transformations.

Week 2, Feb 5- Feb 9

5. Lecture: Linear transformations in geometry, Section 2.2

Rotations, dilations, projections, reflections, rotation-dilations or shears are examples of linear transformations appearing in geometry. How are these transformations described algebraically? A central point is to be able to move back and forth between algebraic and the geometric description. The key fact is that the column vectors v_j of a matrix are the images $v_j = Te_j$ of the basis vectors e_j. We derive for each of the mentioned geometric transformations the matrix form. All of these examples will remain important throughout the course.

6. Lecture: Matrix product and inverse, Section 2.3/4

The composition of linear transformations leads to the **product of matrices**. The **inverse** of a transformation is described by the inverse of the matrix. Square matrices can be treated in a similar way as numbers: we can add them, multiply them with scalars and many matrices have inverses. There is two things to be careful about: the product operation in matrices is not commutative and some nonzero matrices have no inverse. If we take the product of a $n \times p$ matrix with a $p \times m$ matrix, we obtain a $n \times m$ matrix. The dot product as a special case of a matrix product between a $1 \times n$ matrix and a $n \times 1$ matrix. It produces a 1×1 matrix, a scalar. We first look at invertibility of maps $f : X \to X$ in general and then focus on the case of linear maps. If a linear map $\mathbf{R}^n$ to $\mathbf{R}^n$ is invertible, how do we find the inverse? We look at examples when this is the case. Finding x such that $Ax = y$ is equivalent to solving a system of linear equations. Doing it in parallel for all columns gives an elegant algorithm: row reduce the matrix $[A|1_n]$ to end up with $[1_n|A^{-1}]$. We also might have time to see how upper triangular block matrices $\left[\begin{array}{c|c} A & B \\ \hline 0 & C \end{array}\right]$ have the inverse $\left[\begin{array}{c|c} A^{-1} & -A^{-1}BC^{-1} \\ \hline 0 & C^{-1} \end{array}\right]$.

7. Lecture: Image and kernel, Section 3.1

We define the notion of a **linear subspace** of n-dimensional space $\mathbf{R}^n$ and the **span** of a set of vectors. This is a preparation for the more abstract definition of linear spaces which appear later in the course. The main algorithm is the computation of the **kernel** and the **image** of a linear transformation using row reduction. The image of a matrix A is spanned by the columns of A which have a leading 1 in rref(A). The kernel of a matrix A is parametrized by **"free variables"**, the variables for which there is no leading 1 in rref(A). For a $n \times n$ matrix, the kernel is trivial if and only if the matrix is invertible. The kernel is always nontrivial if the $n \times m$ matrix satisfies $m > n$, that is if there are more variables than equations.

Week 3, Feb 12- Feb 16

8. Lecture: Basis and linear independence, Section 3.2

With the previously defined "span" and the newly introduced notion of linear independence, one can define a **basis** for a linear space. A basis is a **set** of vectors which span the space and which are linearly independent. The standard basis in R^n is an example of a basis. We show that if we have a basis, then every vector v can be **uniquely** represented as a linear combination of basis elements $v_1, \dots, v_n$. An extremley important task is to find the basis of the kernel and the basis for the image of a linear transformation.

9. Lecture: Dimension, Section 3.3

In order to be able to define dimension, we need to see why the number of basis elements is independent of the basis. If this is the case, we can define this number to be the **dimension**. The proof uses that if p vectors are linearly independent and q vectors span a linear subspace V, then p is less or equal to q. We see the **rank-nullity theorem**: $\dim\ker(A) + \dim\mathrm{im}(A)$ is the number of columns of A. Even so the result is not difficult to verify, it is also called the **fundamental theorem of linear algebra**. It can e quite useful for us, for example, when understanding why data fitting works.

10. Lecture: Change of coordinates, Section 3.4

Switching to a different basis can be useful when solving problems. For example, to find the matrix of the reflection at a line or projection onto a plane, one can first find the matrix B in a suitable basis $\mathcal{B} = \{v_1, v_2, v_3\}$, then use $A = SBS^{-1}$ to get A. The matrix S contains the basis vectors in the columns $c = S^{-1}v$ are the coordinates in the new basis. In a new basis $Se_i = v_i$ the formula $B = S^{-1}AS$ (BIAS) describes the new matrix.

Week 4, Feb 19-Feb 23

Monday, February 19th, is Presidents day, no class.

11. Lecture: Linear spaces, Section 4.1

The concept of an abstract linear spaces allows to introduce **linear spaces of functions**. This is useful for applications in differential equations. In this lecture we generalize the concept of linear subspaces of $\mathbf{R}^n$ and consider **abstract linear spaces**. An abstract linear space is a set X closed under addition and scalar multiplication and which contains 0. We look at many examples. An important one is the space $X = C([a,b])$ of continuous functions on the interval $[a,b]$ or the space P_5 of polynomials of degree smaller or equal to 5, or the linear space of all 3×3 matrices.

12. Lecture: orthonormal bases and projections, Section 5.1

We review orthogonality between vectors u,v by $u \cdot v = 0$ and define **orthonormal basis**, a basis which consists of unit vectors which are all orthogonal to each other. The orthogonal complement of a linear space V in R^n is defined the set of all vectors perpendicular to all vectors in V. It can be found as a kernel of the matrix which contains a basis of V as rows. We then define orthogonal projection onto a linear subspace V. Given an orthonormal basis $\{u_1, \dots, u_n\}$ in V, we have a formula for the orthogonal projection: $P(x) = (u_1 \cdot x)u_1 + \dots + (u_n \cdot x)u_n$. This simple formula for a projection only holds if we are given an orthonormal basis in the subspace V. We mention already that this formula can be written as $P = AA^T$ where A is the matrix which contains the orthonormal basis as columns.

Week 5, Feb 26-Mar 2

Lecture: Review for midterm on Feb 27

13. Lecture: Gram-Schmidt and QR factorization, Section 5.2

The **Gram Schmidt orthogonalization process** leads to the QR factorization of a matrix A. We will look at this process geometrically as well as algebraically. The geometric process of "straightening out" and "adjusting length" can be illustrated well in 2 and 3 dimensions. Once the formulas for the orthonormal vectors w_j from a given set of vectors v_j are derived, one can rewrite it in matrix form. If the v_j are the m columns of a $n \times m$ matrix A and w_j the columns of a $n \times m$ matrix Q, then $A = QR$, where R is a $m \times m$ matrix. This is the QR factorization. It has its use in numerical methods like for finding the eigenvalues of a matrix.

14. Lecture: Orthogonal transformations, Section 5.3

We first define the **transpose** A^T of a matrix A. **Orthogonal matrices** are defined as matrices for which $A^T A = 1_n$. This is equivalent to the fact that the transformation T defined by A preserves angles and lengths. Rotations and reflections are examples of orthogonal transformations. We point out the difference between orthogonal projections and orthogonal transformations. The identity matrix is the only orthogonal matrix which is also an orthogonal projection. We also stress that the notion of orthogonal matrix only applies to $n \times n$ matrices and that the column vectors form an orthonormal basis. A matrix A for which all columns are orthonormal is not orthogonal if the number or rows is not equal to the number of columns.

Week 6, Mar 5- Mar 9

15. Lecture: Least squares and data fitting, Section 5.4

This is an important lecture from the application point of view. It covers a part of statistics. We learn how to **fit data** with any finite set of functions. To do so, we write the fitting problem as a in general overdetermined system of linear equations $Ax = b$ and find from this the **least square solution** x_* which has geometrically the property that Ax_* is the projection of b onto the image of A. Because this means $A^T(Ax_* - b) = 0$, we get the formula

$$x_* = (A^T A)^{-1} A b \ .$$

An example is to fit a set of data (x_i, y_i) by linear functions $\{f_1, \dots, f_n\}$. This is powerful as we can fit by any type of functions, even functions of several variables.

16. Lecture: Determinants I, Section 6.1

One way to define the determinant of a $n \times n$ matrix is the **permutation definition** which was also historically given first by Leibniz. This setup immediately implies the Laplace expansion formula and allows comfortably to derive all the properties of determinants from the original definition. In this lecture, students learn about permutations in terms of **patterns** which can be used to illustrate permutations and signatures. "Patterns" are permutations and the signature is determined by the "number of **"up-crossings"**. In this lecture, we see the definition of determinants in all dimensions, see how it fits with 2 and 3 dimensional case. We practice already **Laplace expansion** to compute determinants.

17. Lecture: Determinants II, Section 6.2

We learn about the **linearity property** of determinants and how Gauss-Jordan elimination allows a fast computation of determinants. The computation of determinants by Gauss-Jordan elimination is quite efficient. Often we can see the determinant already after a few steps because the matrix has become upper triangular. We also point out how to compute determinants for partitioned matrices. We do lots of examples, also harder examples in which we learn how to decide which of the methods to use: permutation method, Laplace expansion, row reduction to a triangular case or using partitioned matrices.

Spring break: Mar 10- Mar 18

Week 8, Mar 20-Mar 24

18. Lecture: Eigenvalues, Section 7.1-2

Eigenvalues and eigenvectors for square matrices are introduced in this lecture. It is good to see them first in concrete examples like rotations, reflections, shears. As the book, we can motivate the concept using **discrete dynamical systems**, like the problem to find the growth rate of the Fibonacci sequence. It is for such problems, where it becomes evident, why the computation of eigenvalues and eigenvectors is useful.

19. Lecture: Eigenvectors, Section 7.3

The computation of eigenvectors relates to the computation of the kernel of a linear transformation as $Av = \lambda v$ can be rewritten as $(A - \lambda I)v = 0$. We give also a geometric idea what eigenvectors are and look at lots of examples. A good class of examples are **Markov matrices**, which are important from the application point of view. Markov matrices always have an eigenvalue 1 because the transpose has an eigenvector $[1, 1, \dots 1]^T$. The eigenvector of A to the eigenvalue 1 has significance. It belongs to a stationary probability distribution.

20. Lecture: Diagonalization, Section 7.4

A major result of this section is that if all eigenvalues of a matrix are different, one can **diagonalize** the matrix A. There is an **eigen basis**. We also see that if the eigenvalues are the same, like for the shear matrix, one can not diagonalize A. If the eigenvalues are complex like for a rotation, one can not diagonalize over the reals. Since we like to able to diagonalize in as many situations as possible, we allow complex eigenvalues from now on.

Week 8, Mar 26-Mar 30

21. Lecture: Complex eigenvalues, Section 7.5

We start with a short review on complex numbers. Course assistants will do more to get the class up to speed with complex numbers. We take for granted the **fundamental theorem of algebra** which assures that a polynomial of degree n has n solutions, when counted with multiplicities. We express the determinant and trace of a matrix in terms of eigenvalues. Unlike in the real case, these formulas hold for any matrix.

22. Lecture: Stability, Section 7.6

We study the **stability problem** for discrete dynamical systems. The absolute value of the eigenvalues determines the stability of the transformation. If all eigenvalues are in absolute value smaller than 1, then the origin is **asymptotically stable**. A good example to discuss is the case, when the matrix is not diagonalizable, like for example for a shear dilation $S = \begin{bmatrix} 0.99 & 1000 \\ 0 & 0.99 \end{bmatrix}$, where the expansion by the off diagonal shear competes with the contraction in the diagonal.

23. Lecture: Symmetric matrices, Section 8.1

The main point of this lecture is to see that **symmetric matrices** can be diagonalized. This is the spectral theorem. The key fact is that the eigenvectors of a symmetric matrix are perpendicular to each other. An intuitive proof of the spectral theorem can be given in class: after a small perturbation of the matrix all eigenvalues are different and diagonalization is possible. When making the perturbation smaller and smaller, the eigenspaces stay perpendicular and in particular linearly independent. The shear is the prototype of a matrix, which can not be diagonalized. This lecture also gives plenty of opportunity to practice finding an eigenbasis and possibly for Gramm-Schmidt, if an orthonormal eigenbasis needs to be found in a higher dimensional eigenspace.

Week 9, Apr 2 -Apr 6

24. Lecture: Review for second midterm on Apr 3

We review for the second midterm in section. Since a plenary review for all students covers the theory, one can focus on student questions and see the big picture or discuss some True/False problems or practice exam problems.

25. Lecture: Differential equations I, Section 9.1

We learn to solve **systems of linear differential equations** by diagonalization. We discuss linear stability of the origin. Unlike in the discrete time case, where the absolute value of the eigenvalues mattered, the real part of the eigenvalues is now important. We also keep in mind the one-dimensional case, where these facts are obvious. The point is that linear algebra allows to reduce the higher dimensional case to the one-dimensional case.

26. Lecture: Differential equations II, Section 9.2

A second lecture is necessary for the important topic of applying linear algebra to **solve differential equations** $x' = Ax$, where A is a $n \times n$ matrix and $x(t)$ is a time dependent vector. While the central idea is to diagonalize A and solve $y' = By$, where B is diagonal, we can do so faster: write the initial condition $x(0)$ as a linear combination of eigenvectors $x(0) = a_1 v_1 + \ldots + a_n v_n$, then get $x(t) = a_1 v_1 e^{\lambda_1 t} + \ldots + a_n v_n e^{\lambda_n t}$. We also look at examples where the eigenvalues λ_1 of the matrix A are complex. An important case for the later is the **harmonic oscillator** with and without damping. There would be many more interesting examples from physics.

Week 10, Apr 9 - Apr 13

27. Lecture: Nonlinear systems, Section 9.4

This section is covered by a separate handout numbered Section 9.4. How can **nonlinear differential equations** in two dimensions $\dot{x} = f(x, y), \dot{y} = g(x, y)$ be analyzed using linear algebra? The key concepts are finding **null clines**, **equilibria** and their nature using **linearization** of the system near the equilibria by computing the **Jacobean matrix**. Good examples are **competing species systems** like the example of Murray, **predator-pray examples** like the Volterra system or **mechanical systems** like the pendulum.

28. Lecture: Operator Method, Section 4.2

The operator $Df = f'$ as well as polynomials of the operator D allows solve linear higher order differential equations $p(D) = g$ using the **operator method**. The method generalizes the integration process which we use to solve for examples like $f''' = \sin(x)$ where three fold integration leads to the general solution f. For a problem $p(D) = g$ we factor the polynomial $p(D) = (D - a_1)(D - a_2) \dots (D - a_n)$ into linear parts and invert each linear factor $(D - a_i)$ using an integration factor. This operator method is very general and always works. It also provides us with a justification for a more convenient way to find solutions.

29. Lecture: Cookbook method

This operator method to solve differential equations $p(D)f = g$ works unconditionally. It allows also to formulate a "cookbook method". This rule describes, how to find the special solution of the inhomogeneous problem by first finding the general solution to the **homogeneous equation** and then finding a special solution. Very important cases are the situation $\dot{x} - ax = g(x)$, the **driven harmonic oscillator**

$$\ddot{x} + c^2 x = g(x)$$

or the **driven damped harmonic oscillator**

$$\ddot{x} + b\dot{x} + c^2 x = g(x)$$

Special care has to be taken if $g(x)$ is in the kernel of $p(D)$ or if the polynomial p has repeated roots.

Week 11, Apr 16 - Apr 20

30. Lecture: Fourier I

inner products in linear spaces. It generalizes the dot product. For 2π-periodic functions, one takes $\langle f, g \rangle$ as the integral of $f\ddot{g}$ from $-\pi$ to π and divide by 2π. The expansion of a function with respect to the orthonormal basis $1/\sqrt{2}, \cos(nx), \sin(nx)$ leads to the **Fourier expansion**

$$f(x) = a_0(1/\sqrt{2}) + \sum_{n=1}^{\infty} a_n \cos(nx) + b_n \sin(nx) \ .$$

A nice example to see how Fourier theory is useful is to derive the **Leibniz series** for $\pi/4$ by writing

$$x = \sum_{k=1}^{\infty} 2\frac{(-1)^{k+1}}{k} \sin(kx)$$

and evaluate it at $\pi/2$.

31. Lecture: Fourier II

The main motivation is that the Fourier basis is an orthonormal eigen basis to the operator D^2. It diagonalizes this operator because $D^2 \sin(nx) = -n^2 \sin(nx), D^2 \cos(nx) = -n^2 \cos(nx)$. **Parseval's identity**

$$||f||^2 = a_0^2 + \sum_{n=1}^{\infty} a_n^2 + b_n^2 \ .$$

is the "Pythagorean theorem" for function spaces. It is useful to estimate how fast a finite sum converges. We mention also applications like computations of series by the **Parseval's identity** or by relating them to a Fourier series. Nice examples are the derivation of formulas for $\zeta(2)$ or $\zeta(4)$ using the Parseval's identity.

32. Lecture: Heat equation

We will use this to solve partial differential equations in the same way as we solved ordinary differential equations. **Linear partial differential equations** $u_t = p(D)u$ or $u_{tt} = p(D)u$ with a polynomial p are solved in the same way as ordinary differential equations: by diagonalization. Fourier theory achieves that the "matrix" D is diagonalized and so the polynomial $p(D)$. This is much more powerful than the separation of variable method, which we do **not** do in this course.

Week 12, Apr 23 - Apr 25

33. Lecture: Wave equation

For example, the partial differential equation

$$u_{tt} = u_{xx} - u_{xxxx} + 10u$$

can be solved with Fourier in the same way as we solve the wave equation. The method allows even to solve partial differential equations with a driving force like for example

$$u_{tt} = u_{xx} - u + \sin(t) \ .$$

34. Lecture: Summary and overview

More examples both for Fourier and PDEs overview.
Reading Period: April 26 - May 2, Exam Period May 3- May 12.
Oliver Knill, April 4, 2018