JOURNAL OF MATHEMATICS FOR DATA SCIENCE/BIOSTATISTICS TRENTO, A.A. 2019/20

INSTRUCTORS: ANDREA CARANTI, SIMONE UGOLINI

Note. The description of the content of a yet to be delivered lecture is meant to be part of the planning.

Lecture 1. (Caranti) Monday 16 September 2019 (2 hrs)

The statements $0 = 0$ and $0 = 1$. Numbers: natural numbers, integers, rational numbers, real numbers. If $a \neq 0$ is a real number, it has an inverse a^{-1}.

Spurious solutions can be introduced while manipulating an equation.

One linear equation in one unknown. Discussion of the various cases.

One linear equation in two unknowns. Discussion of the various cases.

Two linear equations in two unknowns. The substitution method. The elimination method.

Lecture 2. (Ugolini) Thursday 19 September 2019 (2 hrs)

The elimination method (2 linear equations in 2 unknowns).

Geometrical interpretation of the solutions of a linear system of 2 linear equations in 2 unknowns (no solution, unique solution, infinitely many solutions) as the relative position of two lines in the plane.

Applied vectors in the plane. Equivalent applied vectors. Vectors of $\mathbf{R}^2$. Length of a vector. Multiplication of a vector by a scalar.

Lecture 3. (Ugolini) Monday 23 September 2019 (2 hrs)

Sum of vectors in $\mathbf{R}^2$ and properties.

If $a, b \in \mathbf{R}$ are both non-zero and if x_0, y_0 is a solution of $ax + by = c$, then the solutions of $ax + by = c$ are exactly all the pairs x, y of the form

$$\begin{cases} x = x_0 - bt \\ y = y_0 + at \end{cases}$$

where t is an arbitrary real number.

Parametric equation of a line in the plane.

The set $\mathbf{R}^n$ of the n-tuples of real numbers together with the operations of sum of vectors and multiplication by scalars is a vector space.

Formal definition of vector space.

Date: Trento, A. A. 2019/20.

Lecture 4. (Ugolini) Thursday 26 September 2019 (2 hrs)

Matrices: definition and notations. Some families of matrices: row vectors, column vectors, square, diagonal and identity matrices. Sum of matrices and multiplication by scalars.

The set V of real matrices $m \times n$ is a vector space (sketch of proof).

Product of matrices. If the product AB is defined, then BA is not necessarily defined. When both products are defined, AB can be different from BA.

Some properties of the sum and product of matrices (if the operations are defined):

- $A(B + C) = AB + AC$;
- $(A + B)C = AC + BC$;
- $(AB)C = A(BC)$.

Lecture 5. (Ugolini) Monday 30 September 2019 (2 hrs)

Homogeneous linear systems.

Elementary row operations. Elementary matrices: S_{ij}, $D_i(k)$ and $E_{ij}(k)$. Row echelon form (REF).

Any matrix can be reduced to REF through Gaussian elimination (without proof). The REF of a non-zero matrix is not unique.

Exercise: finding all solutions of a homogeneous linear system.

Lecture 6. (Caranti) Thursday 3 October 2019 (2 hrs)

Geometric interpretation of homogeneous systems of two equations in three unknowns: planes and lines in the space.

Homogeneous systems of three equations in three unknowns: an example where the solutions form a line

Lecture 7. (Ugolini) Monday 7 October 2019 (2 hrs)

Reversibility of elementary row operations.

Vector subspaces (definition).

The set of all the solutions of $Ax = O$ is a vector subspace of $\mathbf{R}^n$. If $v_1, v_2, \dots, v_k$ are k vectors of a vector space V, then the set

$$\{ c_1 v_1 + \dots + c_k v_k : c_1, \dots, c_k \in \mathbf{R} \}$$

of all the linear combinations of the k vectors is a vector subspace of V (alternative notations: span $\{ v_1, \dots, v_k \}$ or $\langle v_1, \dots, v_k \rangle$).

Examples of vector subspaces: lines through the origin, planes through the origin. Parametric equation of a plane in the space.

Lecture 8. (Caranti) Thursday 10 October 2019 (2 hrs)

A line in space: from parametric equations to Cartesian ones.

First method: If the parametric equations are

$$\begin{cases} x_1 = f_1 + g_1 t \\ x_2 = f_2 + g_2 t \\ x_3 = f_3 + g_3 t \end{cases}$$

where t is the parameter, eliminate the parameter t from the first two equations by taking

$$\begin{aligned} g_2 x_1 - g_1 x_2 &= g_2(f_1 + g_1 t) - g_1(f_2 + g_2 t) \\ &= g_2 f_1 - g_1 f_2 + g_2 g_1 t - g_1 g_2 t \\ &= g_2 f_1 - g_1 f_2. \end{aligned}$$

Do the same with the first and second equation, and with the second and third. Any two of the resulting equations will do.

Second method: Find the *unknowns* a, b, c such that $ax_1 + bx_2 + cx_3$ is a constant not containing t. So

$$\begin{aligned} ax_1 + bx_2 + cx_3 &= a(f_1 + g_1 t) + b(f_2 + g_2 t) + c(f_3 + g_3 t) \\ &= af_1 + bf_2 + cf_3 + (ag_1 + bg_2 + cg_3)t. \end{aligned}$$

To let t disappear, we need to solve the equation

$$ag_1 + bg_2 + cg_3 = 0$$

in the unknowns a, b, c.

Testing a subset $\mathcal{S}$ of $\mathbf{R}^3$ for being a subspace. We have only checked whether $O = (0, 0, 0)$ is in $\mathcal{S}$ or not. How to build examples of this kind.

Systems of non-homogenous linear equations. Method for finding solutions (without proof): take the *complete* matrix of coefficients (I have not defined the term in the lecture), do the reduction to REF form, and go back to a system. We have seen that the lack of solutions manifests itself in an equation becoming something like $0 = 3$ (no explanation yet, no mention of pivot in the last column). How to build examples.

Lecture 9. (Ugolini) Monday 14 October 2019 (2 hrs)

Finding the solutions of a linear system of two equations in 3 unknowns (one exercise).

Proposition (with proof). If x_0 is a solution of the linear system $Ax = b$, then the solutions of $Ax = b$ are all the vectors $x_0 + y$, where$y$ is a solution of $Ax = O$.

Proposition (with proof). Let $Ax = b$ be a linear system. Let $[A' \mid b']$ be a REF matrix obtained from $[A \mid b]$ through elementary row operations. If b' contains one pivot, then $Ax = b$ has no solution.

Definitions of matrix in quasi row reduced echelon form (qRREF) and in row reduced echelon form (RREF). Examples.

Definition of system of generators of a vector space: if V is a vector space and $v_1, \ldots, v_n$ are n vectors of V, then we say that $v_1, \ldots, v_n$ generate V if $span\{v_1, \ldots, v_n\} = V$. Examples in $\mathbb{R}^2$.

Transpose of a matrix: definition and one example.

Lecture 10. (Caranti) Thursday 17 October 2019 (2 hrs)

If $v_1, \dots, v_m$ are a system of generators for the vector space V, this means (by the very definition) that given any $v \in V$, there are $a_1, \dots, a_m \in \mathbf{R}$ such that

$$v = a_1 v_1 + \dots + a_m v_m. \tag{1}$$

However, we see on on example that the representation (1) is not necessarily unique, in the sense that we may have

$$v = a_1 v_1 + \dots + a_m v_m = b_1 v_1 + \dots + b_m v_m$$

with $(a_1, \dots, a_m) \neq (b_1, \dots, b_m)$.

Definition of linearly dependent and linearly independent vectors. Thus $v_1, \dots, v_m$ are dependent if the equation

$$a_1 v_1 + \dots + a_m v_m = 0 \tag{2}$$

has a solution different from the trivial solution $a_1 = \dots = a_m = 0$. Whereas they are independent if the only solution to (2) is the trivial one $a_1 = \dots = a_m = 0$. Examples.

Theorem (with proof). The following are equivalent

(1) $v_1, \dots, v_m$ are linearly dependent, and
(2) (at least) one of the v_i can be written as a linear combination of the others.

As a consequence, if $v_1, \dots, v_m$ are a system of generators of a space V, and they linearly dependent, say v_m is a linear combination of $v_1, \dots, v_{m-1}$, then $v_1, \dots, v_{m-1}$ is a system of generators for V, that is, you can safely drop v_m.

Bases: definition.

Theorem (with proof). The following are equivalent

(1) $v_1, \dots, v_m$ are a basis of the space V, and
(2) $v_1, \dots, v_m$ are a system of generators for V, and they are linearly independent.

Lecture 11. (Ugolini) Monday 21 October 2019 (2 hrs)

The REF of a matrix is not unique (example).

Theorem (without proof). The RREF of a matrix is unique.

Relative position of two planes in the space: coincident, parallel and not coincident, incident planes. There are infinitely many pairs of planes intersecting along a line. Example of different Cartesian equations of a line defined through parametric equations. Review of the definition of basis and examples. Standard basis of $\mathbf{R}^n$. Standard basis of the vector space of the matrices $m \times n$. Example of different bases for the same vector space.

Theorem (without proof). If $v_1, \dots, v_m$ and $w_1, \dots, w_n$ are two different bases of a vector space V, then $m = n$.

The number of vectors in a basis of a vector space V is called the dimension of V ($\dim(V)$).

Lecture 12. (Ugolini) Thursday 24 October 2019 (2 hrs)

Let A be a $m \times n$ matrix.

- The *rank* of A ($rank(A)$) is the number of pivots of a REF of A (in particular, $rank(A)$ is the number of pivots of the RREF of A).
- The *nullity* of A is $n - rank(A)$.
- The *nullspace* of A is the space of solutions of $Ax = O$.

The following bounds hold: $0 \leq rank(A) \leq \min\{m, n\}$. Moreover, $rank(A) = 0$ if and only if A is a zero matrix.

Finding a basis for the nullspace.

Theorem (without proof). Let $[I_k|C]$ be the RREF of A. Then a basis of the nullspace of A is given by the columns of the matrix $\left[\frac{-C}{I_{n-k}}\right]$.

Examples. How to find a basis of the nullspace when the RREF has not the form $[I_k|C]$ (elimination of zero-rows and permutation of columns).

Theorem of Rouché-Capelli (without proof). Let $Ax = b$ be a linear system, where A is a $m \times n$ matrix.

- If $rank(A) \neq rank([A|b])$, then the system has no solution.
- If $rank(A) = rank([A|b])$, then the system has solutions. Moreover, if $rank(A) = n$, then the solution is unique, while there are infinitely many solutions if $rank(A) < n$. In this latter case, there are $n - rank(A)$ free variables.

Lecture 13. (Ugolini) Monday 28 October 2019 (2 hrs)

Review exercises on linear systems, systems of generators and bases.

Theorem (with proof). Let V be a vector space with $\dim(V) = n$. Let $b_1, \dots, b_n$ be a basis of V. If $v_1, \dots, v_k$ are linearly independent vectors of V, then they can be extended to a basis of V.

Theorem (without proof). Let V be a vector space with $\dim(V) = n$. If $v_1, \dots, v_n$ are linearly independent vectors, then $v_1, \dots, v_n$ are a basis of V.

Lecture 14. (Caranti) Thursday 31 October 2019 (2 hrs)

Four lemmata on systems of generators, linearly independent vectors and bases.

Some related exercises: extending a linearly independent system of vectors to a base, extracting a base from a system of generators.

Vector space of polynomials (of degree less than n) — computing with the n-tuples of coefficients.

Lecture 15. (Caranti, Ugolini) Monday 4 November 2019 (3 hrs)

Written test

Lecture 16. (Caranti) Thursday 7 November 2019 (2 hrs)

Computing the inverse of a matrix. Determinants from RREF.

Lecture 17. (Ugolini) Monday 11 November 2019 (2 hrs)

How to find a basis of $V = \langle v_1, \dots, v_k \rangle$, where $v_1, \dots, v_k$ are vectors of $\mathbf{R}^n$:

- construct a matrix A having $v_1, \dots, v_k$ as columns;
- reduce A to a REF matrix A';
- the columns of A' containing pivots correspond to the columns of A which form a basis of V.

All the vector spaces of dimension n are isomorphic to $\mathbf{R}^n$.

- The space P_n of real polynomials of degree smaller than n is isomorphic to $\mathbf{R}^n$.
- The space $\mathcal{M}_{m\times n}$ of the real matrices $m \times n$ is isomorphic to $\mathbf{R}^{mn}$.

Computing determinants of matrices.

- If A is a triangular matrix, then $\det(A)$ is equal to the product of the entries on the main diagonal.
- If A is in REF, then A is upper triangular.
- If A is not in REF, then we can compute $\det(A)$ reducing A to a REF A'. The effects of elementary row operations on determinants are the following:
  - when we switch two rows, the determinant is multiplied by -1;
  - when we multiply a row by a number $k \neq 0$, the determinant is multiplied by k;
  - adding a row or a multiple of a row to another row has no effect on the determinant.

Theorem (with proof). Let A be a $n \times n$ matrix.
The system $Ax = b$ has a unique solution if and only if $\det(A) \neq 0$.

Lecture 18. (Caranti) Thursday 14 November 2019 (2 hrs)

Calculating determinants: Laplace expansion, Sarrus rule.
Linear maps, bases and matrices.

Lecture 19. (Ugolini) Monday 18 November 2019 (2 hrs)

Let V and W be two vector spaces with $\dim(V) = n$ and $\dim(W) = m$. If $b_1, \dots, b_n$ are a basis of V and $g_1, \dots, g_m$ are a basis of W, then the matrix A of f with respect to such bases is $A = [f(b_1)| \dots |f(b_n)]$ where the j-th column contains the coordinates of $f(b_j)$ with respect to the basis $g_1, \dots, g_m$.

If $b_1, \dots, b_n$ and $b'_1, \dots, b'_n$ are two bases of a vector space V, then there exists a matrix M (matrix of change of bases) such that $B' = BM$, where $B = [b_1| \dots |b_n]$ and $B' = [b'_1| \dots |b'_n]$. Moreover, if X is the vector of the coordinates of a vector $v \in V$ with respect to the basis B and X' is the vector of the coordinates with respect to the basis B', then $X = MX'$.

Let A be the matrix of a linear function $f : V \to V$ with respect to a basis $E = [e_1| \dots |e_n]$. Then $Y = AX$, where X and Y are the coordinates of the input and output vectors with respect to E. If $E' = [e'_1| \dots |e'_n]$ is another basis of V and $E' = EM$, then $Y' = M^{-1}AMX'$, where X' and Y' are the coordinates of the input and the output vectors with respect to the basis E'. Then $B = M^{-1}AM$ is the matrix of f with respect to the basis E'.

Lecture 20. (Caranti) Thursday 21 November 2019 (2 hrs)

Eigenvalues and eigenvectors for a matrix. The eigenvalues of a matrix B are the roots of the characteristic polynomial $\det(B - \lambda I)$. Diagonalisable matrices.The matrix

$$\begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}.$$

Lecture 21. (Ugolini) Monday 25 November 2019 (2 hrs)

Let A be a $n \times n$ matrix. If $\lambda \in \mathbf{R}$ is an eigenvalue of A, then the vector space $E_\lambda(A)$ generated by all the eigenvectors associated with λ is called the eigenspace of λ.

Let $f : V \to V$ be a linear function. A non-zero vector $v \in V$ is an eigenvector of f is there exists a $\lambda \in \mathbf{R}$ such that $f(v) = \lambda v$.

Theorem (with proof). Let A and B be the matrices of the linear function $f : V \to V$ with respect to different bases of V. If

$$\begin{aligned} p(\lambda) &= \det(A - \lambda I), \\ q(\lambda) &= \det(B - \lambda I), \end{aligned}$$

then $p(\lambda) = q(\lambda)$.

We recalled that a $n \times n$ matrix is diagonalisable if and only if there exists a basis of $\mathbf{R}^n$ formed by eigenvectors $v_1, \dots, v_n$ of A. Moreover, if we construct the matrix $S = [v_1|\dots|v_n]$ having $v_1, \dots v_n$ as columns, then

$$S^{-1}AS = D = \begin{bmatrix} \lambda_1 & 0 & \dots & \dots \\ 0 & \lambda_2 & \dots & \dots \\ \dots & \dots & \dots & \\ \dots & \dots & \dots & \lambda_n \end{bmatrix},$$

namely D is the diagonal matrix whose entry $d_{ii} = \lambda_i$, where λ_i is the eigenvalue of v_i.

The following matrix is diagonalisable:

$$\begin{bmatrix} -4 & -3 \\ 10 & 7 \end{bmatrix}.$$

The following matrix is not diagonalisable:

$$\begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}.$$

The following matrix is diagonalisable:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Lecture 22. (Caranti) Thursday 28 November 2019 (2 hrs)

Algebraic and geometric multiplicity of an eigenvalue. Relation to diagonalisability.

Eigenvectors related to distinct eigenvalues are independent (proof only for two).

Examples.

Lecture 23. (Ugolini) Monday 2 December 2019 (2 hrs)

Euclidean norm $\|v\|$ of a vector $v \in \mathbf{R}^n$ and its properties. Unit vectors. Standard scalar product (dot product) $u \cdot v$ in $\mathbf{R}^n$.

Properties. If $u, v, w \in \mathbf{R}^n$ and $c \in \mathbf{R}$, then

- $u \cdot v = v \cdot u$;
- $(u + v) \cdot w = u \cdot w + v \cdot w$;
- $u \cdot (v + w) = u \cdot v + u \cdot w$;
- $(cu) \cdot v = c(u \cdot v)$.

Pythagoras' theorem and its reverse.

Theorem. Let v, w be vectors in $\mathbf{R}^2$ or $\mathbf{R}^3$.

Then $\|v\|^2 + \|w\|^2 = \|v - w\|^2$ if and only if $v \perp w$.

By definition, we say that two vectors $v, w \in \mathbf{R}^n$ are orthogonal if $v \cdot w = 0$.

If a is a non-zero vector in $\mathbf{R}^n$ and $v \in \mathbf{R}^n$, then the projection of v onto a is the vector w, where

$$w = proj_a(v) = \frac{v \cdot a}{a \cdot a} a.$$

A basis $b_1, \dots, b_n$ of $\mathbf{R}^n$ is orthogonal if $b_j \cdot b_j = 0$ for $i \neq j$. Moreover, if $\|b_i\| = 1$ for all i, the basis is orthonormal.

We say that k vectors $v_1, \dots, v_k$ in $\mathbf{R}^n$ are mutually orthogonal if $v_i \cdot v_j = 0$ for $i \neq j$.

Theorem (with proof). If $v_1, \dots, v_k$ are mutually orthogonal non-zero vectors in $\mathbf{R}^n$, then they are linearly independent.

Corollary. If $v_1, \dots, v_n$ are n mutually orthogonal non-zero vectors in $\mathbf{R}^n$, then they are a basis of $\mathbf{R}^n$.

Lecture 24. (Ugolini) Thursday 5 December 2019 (2 hrs)

Theorem (with proof). Let $b_1, \dots, b_n$ be an orthonormal basis of $\mathbf{R}^n$. If $v \in \mathbf{R}^n$, then $v = (v \cdot b_1)b_1 + (v \cdot b_2)b_2 + \cdots + (v \cdot b_n)b_n$.

Gram-Schmidt orthonormalization process. Orthogonal matrices, transpose of a matrix. If A, B are two matrices, then $(AB)^T = B^T A^T$. Symmetric matrices.

Theorem [spectral theorem] (without proof). If A is a symmetric real matrix, then all the eigenvalues of A are real and there exists an orthogonal matrix S such that $S^T A S = D$, where D is a diagonal matrix whose diagonal entries are the eigenvalues of A.

Lecture 25. (Ugolini) Monday 9 December 2019 (2 hrs)

Theorem (with proof). Let A be a symmetric matrix. Let v be an eigenvector of A with eigenvalue λ. Let w be an eigenvector of A with eigenvalue μ. If $\lambda \neq \mu$, then $v \cdot w = 0$.

Affine frames and associated affine coordinates.

Principal component analysis. If C is the covariance matrix of a set of m-dimensional samples with sample mean $\overline{X}$, then C can be orthogonally diagonalized through an orthogonal matrix

$$S = [s_1 | s_2 | \cdots | s_m]$$

where $s_1, s_2, \dots, s_m$ are the columns of S. Such columns are m principal directions of C.

If the column vector $v \in \mathbf{R}^m$ is a (new) sample, then the coordinates of v with respect to the affine frame $(\overline{X}, s_1, \dots, s_m)$ are called the principal components of v.

Dipartimento di Matematica, Università degli Studi di Trento, via Sommarive 14, 38123 Trento

Email address: andrea.caranti@unitn.it

URL: http://www.science.unitn.it/∼caranti/