Complex Numbers

Definition imaginary unit: $i = \sqrt{-1}$ is the imaginary unit
Definition complex number: $z = x + yi$ with $x, y \in \mathbb{R}$
Definition complex numbers: $\mathbb{C} = \{x + yi | x, y \in \mathbb{R}\}$
Definition imaginary/real part: $Im(z) = y$ and $Re(z) = x$
Complex numbers are drawn in the complex plane. The above described form is called normal form.

polar form

Complex numbers can be described by the polar coordinates.
Definition: radius/distance $r := |z| = \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}$
Definition: angle/argument $\theta := Arg(z) = \angle$x-axis and vector
We have $z = x + iy = r(\cos(\theta) + i\sin(\theta))$.
Definition: Euler formula $e^{i\theta} = \cos\theta + i\sin\theta$
Definition: polar representation $z = re^{i\theta}$
The polar representation is not unique. Therefore, often $\theta \in [0; 2\pi]$. Remember: $e^{i\pi} + 1 = 0$ and $e^{2i\pi} = 1$.

polar and normal form conversion

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin\theta$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\cos\theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{1}}{2}$	0
$\tan\theta$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	∞

polar form $\rightarrow$ normal form
$x = r\cos\theta$ and $y = r\sin\theta$
normal form $\rightarrow$ polar form
$r = \sqrt{x^2 + y^2}$.
For x, y one needs $\cos\theta = \frac{x}{r}, \sin\theta = \frac{y}{r}, \tan\theta = \frac{y}{x}$.
Use above table to get θ or compute arctan with:

$$\theta = \begin{cases} \arctan\frac{y}{x} & \text{falls } z \text{ im 1. Quadranten oder auf der positiven } x\text{-Achse liegt} \\ \frac{\pi}{2} & \text{falls } x = 0 \text{ und } y > 0 \\ \pi + \arctan\frac{y}{x} & \text{falls } z \text{ im 2. oder 3. Quadranten liegt} \\ \frac{3\pi}{2} & \text{falls } x = 0 \text{ und } y < 0 \\ 2\pi + \arctan\frac{y}{x} & \text{falls } z \text{ im 4. Quadranten liegt} \end{cases}$$

calculating with $\mathbb{C}$

$z = x + iy = re^{i\theta}$ and $w = z + iv = be^{i\alpha}$ and $\alpha \in \mathbb{R}$
Addition
$z + w := (x + z) + (y + v)i$
Multiplication
$\alpha \cdot z := \alpha x + \alpha yi = \alpha re^{i\theta}$
$w \cdot z := ux - vy + i(vx + uy) =, z \cdot w := re^{i\theta}be^{i\alpha} = rbe^{i(\theta+\alpha)}$

complex conjugate
$\overline{z} := x - iy$ and notice that $z\overline{z} = x^2 + y^2 \geq 0$.
absolute value
$|z| := \sqrt{z\overline{z}}$
division
normal form: $\frac{z}{w} = \frac{z}{w} \cdot \frac{\overline{w}}{\overline{w}} = \frac{z\overline{w}}{|w|^2}$ with $\tau = \frac{1}{|w|^2}$.
polar form: $\frac{z}{w} = \frac{r}{s}e^{i(\theta-\alpha)}$ - $w \neq 0$ of course!
further computation rules

- $\overline{zw} = \overline{z} \cdot \overline{w}$
- $\overline{\frac{z}{w}} = \frac{\overline{z}}{\overline{w}}$
- $\overline{\overline{z}} = z$
- $|z| = |\overline{z}|$
- $|zw| = |z| \cdot |w|$
- $|\frac{z}{w}| = \frac{|z|}{|w|}$
- $|z + w| \leq |z| + |w|$ (triangle inequality)

subsets of $\mathbb{C}$ in the complex plane
$|z - i| = 1$ is the unit circle around i.
potentiation
$z^x = r^x e^{xi\theta}$
roots
$a \in \mathbb{C}$, $n \in \mathbb{N}$. If $n \neq 0$, there are n roots of a. n-th roots of a are $z \in \mathbb{C}$ with $z^n = a$. Thus, $z^n = r^n e^{in\theta} = a = se^{i\alpha}$.
We get $r = \sqrt[n]{s}$. $n\theta = \alpha + 2\pi k \Rightarrow \theta = \frac{\alpha+2\pi k}{n}, k \in \mathbb{Z}$.

Systems of Linear Equations (SLEs)

Definition: SLE A SLE with m linear equations in n unkowns has coefficients a_{ij} $(1 \leq i \leq m, 1 \leq j \leq n)$ and the right-hand side b_i $(1 \leq j \leq m)$ and the unkowns $x_1, ..., x_n$. If $m = n$ the system is square.
A SLE can be written in elimination scheme or with the coefficient matrix $\mathbf{A}$, unknown vector $\mathbf{x}$, and right-hand side $\mathbf{b}$: $\mathbf{Ax} = \mathbf{b}$.
Definition: solution A solution of a SLE is a n-tupel validating all equations. The general solution is the set of all solutions. If a SLE has no solution, it is called inconsistent/unsolvable. If there is exactly one solutoin, it is called uniquely solvable. If there is more than one solution, it is called ambiguously solvable.
Definition: equivalence SLEs with same solutions are equivalent.
Definition: homogenous system A SLE with 0 right-hand side.

Gaussian Elimination

Idea: Transform a SLE in an equivalent but easier to solvable system.
forward elimination
We transform the SLE with elementary row operations to row echelon form, which is easy to solve with back substitution.
Definition: elementary row operations (i) swithcing rows, (ii) adding multiple of rows to other rows, and not necessarily (iii) multiplying a row with a non-zero real number.
Definition: row echelon form For the first non-zero element in each row (called pivot element), all elements below that must be zero and all rows above must have such an element left from the current column. Upper triangular form is a special case for certain $n = m$ matrices.
back substitution
One identifies the last pivot varialbe and substitutes it to the previous equation. From the second last equation one identifies the second last pirvot variable and substitutes it to the previos equation. And so on. Free variables (columns without pivots) are assigned a variable.
procedure (general case)
No deails here, obvious. If $m > r$, the conditions for a solution $c_{r+1} = ... = c_m = 0$ are called consistency conditions. In a homogenous system, consistency conditions are always met.

solution set of a SLE

Definition: rank rank of $\mathbf{A}$ is number of pivot elements r
Theorem: 1.1 A SLE in row echelon form has at least one solution if: $r = m$ or $r < m$ and consistency conditions met.
If solutions exist: unique if $r = n$, $(n - r)$-parameter based if $r < n$.
Corollary: 1.2 The rank r only depends on the coefficient matrix $\mathbf{A}$ but not on the chosen pivot elements or the right-hand side $\mathbf{b}$.
The solution $x_1 = ... = x_n = 0$ is called the trivial solution.
Corollary: An homogenous SLE always has the trivial solution. If $r < n$ it has non-trivial ones too.
1.5
Corollary: 1.6 A squared SLE is solvable for any right-hand side if and only if the homogenous system only has the trivial solution.
Definition: regular/singular If a SLE has a unique solution, it is regular/non-singular. Otherwise it is singular.
$\mathbf{Ax} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combinatin of column vectors of $\mathbf{A}$.

Matrices and Vectors in $\mathbb{R}^n$ and $\mathbb{C}^n$

Matrices, row-/column vectors

Definition: A $m \times n$ matrix $\mathbf{A}$ is a rectangular scheme of mn elements in m rows and n columns. The element in row i, column j is $a_{ij} = (\mathbf{A})_{ij}$. $\mathbf{A} = (a_{ij})$.
Definition: square matrices $n \times n$ matrix - with order n
Definition: null/zero matrix $a_{ij} = 0$ - denoted $\mathbf{O}$
Definition: diagonal elements a_{jj} $(j = 1, ..., min(n, m))$ are diagonal elements. Their set is the (main) diagonal of $\mathbf{A}$.
Definition: diagonal matrix $(\mathbf{A})_{ij} = 0$, $i \neq j$ with $\mathbf{D} = diag(d_{11}, ..., d_{nn})$
Definition: unit matrix/identity $\mathbf{I}_n = diag(1, 1, ..., 1)$
Definition: upper triangular matrix $(\mathbf{R})_{ij} = 0, i > j$
Definition: lower triangular matrix $(\mathbf{L})_{ij} = 0, i < j$
Definition: vector $m \times 1$: column-vector & $1 \times n$: row vector
The set real/complex $m \times n$ matrices is $\mathbb{R}^{m \times n}/\mathbb{C}^{m \times n}$.

calculating with matrices

A a $m \times n$ matrix & $\mathbf{B}$ a $m \times n$ matrix & α a scalar
Definition: scalar multiplication $(\alpha\mathbf{A})_{ij} :\equiv \alpha(\mathbf{A})_{ij}$, $1 \leq i \leq m, 1 \leq j \leq n$
Definition: addition $(\mathbf{A} + \mathbf{B})_{ij} :\equiv (\mathbf{A})_{ij} + (\mathbf{B})_{ij}$, $1 \leq i \leq m, 1 \leq j \leq n$ (only for same-sized matrices!)
Now, $\mathbf{A}$ a $m \times n$ matrix & $\mathbf{B}$ a $n \times p$ matrix
Definition: multiplication $(\mathbf{AB})_{ij} :\equiv \sum_{k=1}^{n}(\mathbf{A})_{ik}(\mathbf{B})_{kj}$ - the dimension of the product is $m \times p$. (only for suitable-sized matrices!)
Theorem: 2.1

- $(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A})$
- $(\alpha\mathbf{A})\mathbf{B} = \alpha(\mathbf{AB}) = \mathbf{A}(\alpha\mathbf{B})$
- $(\alpha + \beta)\mathbf{A} = (\alpha\mathbf{A}) + (\beta\mathbf{A})$
- $\alpha(\mathbf{A} + \mathbf{B}) = (\alpha\mathbf{A}) + (\alpha\mathbf{B})$
- $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
- $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$
- $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- $(\mathbf{A} + \mathbf{B})\mathbf{C} = (\mathbf{AC}) + (\mathbf{BC})$
- $\mathbf{A}(\mathbf{B} + \mathbf{C}) = (\mathbf{AB}) + (\mathbf{AC})\mathbf{A}(\mathbf{B} + \mathbf{C}) = (\mathbf{AB}) + (\mathbf{AC})$

Theorem: 2.2 neutral matrix exists, inverse matrix exists, 'difference' matrix exists
Theorem: 2.3/.4 $\mathbf{A}_{n \times m} = (\mathbf{a}_1 \ \ ... \ \ \mathbf{a}_n)$ and $\mathbf{x}_{n \times 1}$. $\mathbf{Ax} = \mathbf{a}_1 x_1 + ... + \mathbf{a}_n x_n$. $\mathbf{Ae}_j = \mathbf{a}_j$. With $\mathbf{B}_{m \times p} = (\mathbf{b}_1 \ \ ... \ \ \mathbf{b}_p)$: $\mathbf{AB} = (\mathbf{Ab}_1 \ | \ ... \ | \ \mathbf{Ab}_p)$.
with addition: commutative group & with multiplication: non-commutative ring with identity
Definition: linear combination A linear combination of $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ is $\alpha_1\mathbf{a}_1 + ... + \alpha_n\mathbf{a}_n$ with $\alpha_1, ..., \alpha_n$ as scalars.

symetric/hermitian matrices & transpose

Definition: $\mathbf{A}_{n \times m}$. $\mathbf{A}^\top_{m \times n}$ with $(\mathbf{A}^\top)_{ij} :\equiv (\mathbf{A})_{ji}$ is called transpose. $\overline{\mathbf{A}}_{m \times n}$ with $(\overline{\mathbf{A}})_{ij} :\equiv \overline{(\mathbf{A})_{ij}}$ is called complex conjugate for complex $\mathbf{A}$. $\mathbf{A}^H :\equiv (\overline{\mathbf{A}})^\top = \overline{\mathbf{A}^\top}$ is called conjugate/hermitian transpose.
Definition: symmetry $\mathbf{A}$ is symmetric if $\mathbf{A}^\top = \mathbf{A}$. We say it is skew-symmetric if $\mathbf{A}^\top = -\mathbf{A}$.
Definition: hermitian $\mathbf{A}$ is hermitian if $\mathbf{A}^H = \mathbf{A}$.
Theorem: 2.6 $(\mathbf{A}^H)^H = \mathbf{A}$ & $(\alpha\mathbf{A})^H = \overline{\alpha}\mathbf{A}^H$ & $(\mathbf{A} + \mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H$ & $(\mathbf{AB})^H = \mathbf{B}^H\mathbf{A}^H$.
Theorem: 2.7 $\mathbf{A}, \mathbf{B}$ (square) symmetric: $\mathbf{AB} = \mathbf{BA}$ & $\mathbf{A}^H\mathbf{A}$ and $\mathbf{AA}^H$ are symmetric (for arbitrary $\mathbf{A}$)
Corollary: 2.8 $\mathbf{A}_{m \times n}, \mathbf{B}_{n \times p}, \underline{\mathbf{y}} = d(y_1 \ ... \ y_n)$: $\underline{\mathbf{y}}\mathbf{B} = y_1\underline{\mathbf{b}_1} + ... + y_n\underline{\mathbf{b}_n}$ - $e_i^\top\mathbf{B} = \underline{\mathbf{b}_i}$. And $\mathbf{AB} = \begin{pmatrix} \underline{\mathbf{a}_1}\mathbf{B} \\ ... \\ \underline{\mathbf{a}_m}\mathbf{B} \end{pmatrix}$

scalar product, norm, lenth, angles

See at later/generalized section. Chapter 2.4 just instanciation.

outer product, orthogonal projections on a line

See at later/generalized section. Chapter 2.4 just instanciation.

matrices as linear maps

$\mathbf{A}_{m \times n}$ defines map: $\mathbf{A} : \mathbb{E}^n \rightarrow \mathbb{E}^m, \mathbf{x} \mapsto \mathbf{Ax}$. We have characteristics: $\mathbf{A}(\gamma\mathbf{x} + \tilde{\mathbf{x}}) = \gamma(\mathbf{Ax}) + (\mathbf{A}\tilde{\mathbf{x}})$.
Definition: The above is characteristic for linear maps. The domain is $\mathbb{E}^n$. The codomain is $\mathbb{E}^m$.
Definition: image $im \ \mathbf{A} :\equiv \{\mathbf{Ax} \in \mathbb{E}^m; \mathbf{x} \in \mathbb{E}^n\}$
Definition: kernel $ker \ \mathbf{A} := \{\mathbf{x} \in \mathbb{E}^n | \mathbf{Ax} = \mathbf{0}\}$

inverse of a matrix

Definition: ivertibility $\mathbf{A}_{n \times n}$ is invertible if some $\mathbf{X}_{n \times n}$ exists so that $\mathbf{AX} = \mathbf{I} = \mathbf{XA}$. $\mathbf{X}$ is denoted $\mathbf{A}^{-1}$ as inverse of $\mathbf{A}$.
Theorem: 2.16 $\mathbf{A}$ invertible $\rightarrow \mathbf{A}^{-1}$ is unique
Theorem: 2.17 $\mathbf{A}_{n \times n}$: $\mathbf{A}$ is invertible if $rank \ A = n$ & more

Theorem: 2.18 $\mathbf{A}_{n\times n}$ and $\mathbf{B}_{n\times n}$ regular: $\mathbf{A}^{-1}$ regular and $(\mathbf{A}^{-1})^{-1} = A$. $\mathbf{AB}$ is regular and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. $\mathbf{A}^H$ regular and $(\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H$.
We may compute $\mathbf{A}^{-1}$ if exists by solving $\mathbf{AX} = \mathbf{I}$. Instead of $\mathbf{X}$ and $\mathbf{I}$ we may also choose the identity column vectors and $\mathbf{x}$.

orthogonal/unitary matrices

Definition: unitary/orthogonal $\mathbf{A}_{n\times n}$ is unitary if $\mathbf{A}^H\mathbf{A} = \mathbf{I}_n$. It is orthogonal if $\mathbf{A}^\top\mathbf{A} = \mathbf{I}_n$.
Notice, for $\mathbf{A}_{n\times m}$, we only have $\mathbf{A}^H\mathbf{A} = \mathbf{I}$ for unitary columns. We only have $\mathbf{A}\mathbf{A}^H = \mathbf{I}$ if $n = m$.
Definition: $\mathbf{A} \in \mathbb{E}^{n\times n}$ is normal $\Leftrightarrow \mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H$.
Theorem: 2.20 $\mathbf{A}, \mathbf{B} \in \mathbb{E}^{n\times n}$ unitary/orthogonal: $\mathbf{A}$ regular and $\mathbf{A}^{-1} = \mathbf{A}^H$. $\mathbf{A}\mathbf{A}^H = \mathbf{I}_n$. $\mathbf{A}^{-1}$ is unitary. $\mathbf{AB}$ is unitary.
Lemma: $\mathbf{A} \in \mathbb{E}^{m\times n}$. $\mathbf{A}^H\mathbf{A} \in \mathbb{E}^{n\times n}$ is hermititan and positiv semidefinit.
Theorem: 2.21 $\mathbf{A}_{n\times n}$ unitary/orthogonal: linear map of $\mathbf{A}$ is length preserving (isometric) and angle preserving: $\|Ax\| = \|x\|, \langle Ax, Ay\rangle = \langle x, y\rangle$.

structured matrices

Apparently not relevant. (Only in script.)

LU decomposition

Gauss elimination as LU decomposition

Regular $\mathbf{A}$. Gauss elimination $\rightsquigarrow$ $\mathbf{U}$ (upper triangular matrix - row echelon form). For $\mathbf{L}$ we have the coefficients during Gauss elimination (only subtracting rows). The diagonal of $\mathbf{L}$ has only 1s. We have $\mathbf{A} = \mathbf{LU}$ - LU decomposition.
Row changes necessary: $\mathbf{P}$ corresponding permutation matrix. Then: $\mathbf{PA} = \mathbf{LU}$. If not row changes: $\mathbf{P} = \mathbf{I}$.
Theorem: 3.1 Square SLE $\mathbf{Ax} = \mathbf{b}$ with regular $\mathbf{A}$: $\mathbf{PA} = \mathbf{LR}, \mathbf{Lc} = \mathbf{Pb}, \mathbf{Rx} = \mathbf{c}$. Then, $\mathbf{x}$ for some $\mathbf{A}$ may be easily computed for various $\mathbf{b}$.
Solving for some $\mathbf{x}$: $\mathbf{Ax} = \mathbf{b} \Rightarrow \mathbf{PAx} = \mathbf{Pb} \Rightarrow \mathbf{LUx} = \mathbf{Pb}$. Solve for $\mathbf{c}$ first: $\mathbf{L}$=$\mathbf{Pb}$ (forward substitution). Solve for $\mathbf{x}$ second: $\mathbf{Ux} = \mathbf{c}$ (backward substitution).

general case

Theorem: 3.3 $\mathbf{Ax} = \mathbf{b}$ a $m\times n$ SLE: $\mathbf{P}_{m\times m}, \mathbf{R}_{m\times n}, \mathbf{L}_{m\times m}$.
Corollary: 3.4 $\mathbf{A}_{m\times n}$ with $rank\ \mathbf{A} = r$: $\tilde{\mathbf{L}}_{m\times r}$ and $\tilde{\mathbf{U}}_{r\times n}$ and $\mathbf{P}_{m\times m}$ with $\tilde{\mathbf{L}}\tilde{\mathbf{U}} = \mathbf{PA}$.

block LU decomposition & LU updating

Not done.

Cholesky decomposition

Not done.

vector spaces

definition

Definition: Vectorspace (aslo linear space) V over $\mathbb{E}(:\equiv \mathbb{R}$ or $\mathbb{C})$: not empty set with addition $x, y \in V \mapsto x + y \in V$ (inner operation) and multiplication $\alpha \in \mathbb{E}, x \in V \mapsto \alpha x \in V$ (outer operation) as well as eight axioms:

1. $x + y = y + x\ (\forall x, y \in V)$
2. $(x + y) + z = x + (y + z)\ (\forall x, y, z \in V)$
3. $\exists o \in V : x + o = x\ (\forall x \in V)$
4. $\forall x \in V \exists -x \in V : x + (-x) = o$
5. $\alpha(x + y) = \alpha x + \alpha y\ (\forall\alpha \in \mathbb{E}, \forall x, y \in V)$
6. $(\alpha + \beta)x = \alpha x + \beta x\ (\forall\alpha \in \mathbb{E}, \forall x, y \in V)$
7. $(\alpha\beta)x = \alpha(\beta x)\ (\forall\alpha \in \mathbb{E}, \forall x, y \in V)$
8. $1x = x\ (\forall x \in V)$

commutative group regarding addition i(first four axioms)
Elements of V: vectors. $o \in V$: zero vector. Elements of $\mathbb{E}$: scalars.
For $\mathbb{E} = \mathbb{R}$: real vector space. For $\mathbb{E} = \mathbb{C}$: complex vector space.
Theorem 4.1: V over $\mathbb{E}$. $\forall x, y \in V$ and $\forall\alpha \in \mathbb{E}$: $0x = o, \alpha o = o, \alpha x = o \Rightarrow \alpha = 0 \vee x = 0, (-\alpha)x = \alpha(-x) = -(\alpha x)$.
Theorem 4.2: $\forall x, y \in V \exists z \in V : x + z = y$ (z unique)
Definition subtraction: V. $\forall x, y \in V : y - x :\equiv y + (-x)$.
Among others, those are vector spaces: vectors, matrices, continuous real functions, polynomials, real sequences

subspaces, spanning systems

Definition: $\varnothing \neq U \subseteq V$ subspace of V if closed under addition and scalar multiplication: $x + y \in U, \alpha x \in U$ $(\forall x, y \in U, \forall\alpha \in \mathbb{E})$.
Theorem 4.3: A subspace is a vector space itself.
$U_1, ..., U_n$ subspaces over $\mathbb{K}$. Then, $U_1 + ... + U_2 := \{\alpha_1 x_1 + ... + \alpha_n x_n | \alpha_i \in \mathbb{R}, x_i \in U_i\}$ is a vector space.
Theorem 4.4: $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathcal{L}_0$ solution set for $\mathbf{x} \in \mathbb{R}^n$ for $\mathbf{Ax} = o$: $\mathcal{L}_0$ subspace of $\mathbb{R}^n$.
Definition linear combination: V over $\mathbb{E}$. $a_1, ..., a_l \in V$.
$x :\equiv \gamma_1 a_1 + ... + \gamma_l a_l = \sum_{k=1}^{l} \gamma_k a_k$ with $\gamma_1, ..., \gamma_l \in \mathbb{E}$ is linear combination of $a_1, ..., a_l$.
Definition span: Set of all linear combinations of $a_1, ..., a_l$ is the subspace spanned by $a_1, ..., a_l$:
$span\ \{a_1, ..., a_l\} :\equiv \{\sum_{k=1}^{l} \gamma_k a_k; \gamma_1, ..., \gamma_l \in \mathbb{E}\}$
Infinite sequence S or $S \subset V$:
$span\ \ S :\equiv \{\sum_{k=1}^{m} \gamma_k a_k; m \in \mathbb{N}; a_1, ..., a_m \in S; \gamma_1, ..., \gamma_m \in \mathbb{E}\}$
$a_1, ..., a_l$ or S is called spaning set of the span.

linear dependence, bases, dimensions

Definition linear dependence: $a_1, ..., a_l \in V$ linearly dependent if $\gamma_1, ..., \gamma_l \in \mathbb{E}$ not all zero: $\gamma_1 a_1 + ... + \gamma_l a_l = o$. But if $\gamma_1 = ... = \gamma_l = 0$, then linearly dependent.
Lemma 4.5: $l \geq 2$: $a_1, ..., a_l$ linearly dependent $\Leftrightarrow$ one vector linear combination of others
Definition linear dpendence: Infinite set of vectors linearly independent $\Leftrightarrow$ each subset linearly independent.
Definition basis: Linearly independent spanning set of V is called basis of V.
n columns $b_1, ..., b_n \in \mathbb{E}^n$ of $\mathbf{B}_{n\times n}$ basis of $\mathbb{E}^n \Leftrightarrow \mathbf{B}$ regular.
NOW: VECTOR SPACES WITH FINITE SPANNING SETS

Lemma 4.6: Finite spanning set of non-trivial $V \Leftrightarrow$ basis as subset of spanning set exist
Lemma: 4.7 V with finite spanning set: all bases of V same number of vectors.
Definition: Number of basis vectors (in each basis) of V with finite spanning set is called dimension of V: $dim\ V$. Such V is finite-dimensional. For $dim\ V = n$, V is called n-dimensional.
Lemma: 4.8 If $\{b_1, ..., b_m\}$ spans V: Every set $\{a_1, ..., a_l\} \subset V$ of $l > m$ vectors is linearly dependent.
Lemma: 4.9/.11 Every set of linearly independent vectors from V with finite spanning set can be extended to a basis of V.
Corollary: 4.10 V with finite dimension: All sets of $n = dim\ V$ linearly independent vectors are a basis of V.
Theorem: 4.12 $\{b_1, ..., b_n\} \subset V$ is basis of V if and only if $\forall x \in V$ $x = \sum_{k=1}^{n} \zeta_k b_k$ for unique ζ_k.
Definition: cooridnates ζ_k for x are called coordinates of x regarding basis $\{b_1, ..., b_n\}$. $\zeta = (\zeta_1 \ \ \dots \ \ \zeta_n)^\top$ is called coordinate vector. $x = \sum_{k=1}^{n} \zeta_k b_k$ is called representation in coordinates of x.
Definition: complementary subspaces Subspaces U and U' of V with $\forall x \in V : x = u + u'$ with $u \in U, u' \in U'$ are called complementary (subspaces of V). V then is direct sum: $V = U \oplus U'$.

basis change, coordinate transformation

$\{b_1, ..., b_n\}$ 'old' basis of V. $\{b_1', ..., b_n'\}$ 'new' basis of V.
Definition: $b_k' = \sum_{i=1}^{k} \tau_{ik} b_i, k = 1, ..., n$. $\mathbf{T}_{n\times n} = (\tau_{ik})$ is transformation matrix of change of basis.
In k-th column of $\mathbf{T}$: coordinates of k-th new basisvector regarding 'old' basis.
Theorem: 4.13 $x \in V$. $\zeta = (\zeta_1 \ \ \dots \ \ \zeta_n)^\top$ cooridnate vector 'old' basis. $\zeta' = (\zeta_1' \ \ \dots \ \ \zeta_n')^\top$ cooridnate vector 'new' basis. $\sum_{i=1}^{n} \zeta_i b_i = x = \sum_{k=1}^{n} \zeta_k' b_k'$.
Then: $\zeta_i = \sum_{k=1}^{n} \tau_{ik}\zeta_k'$ or rather $\zeta = \mathbf{T}\zeta'$. Because $\mathbf{T}$ regular: $\zeta' = \mathbf{T}^{-1}\zeta$.

linear maps

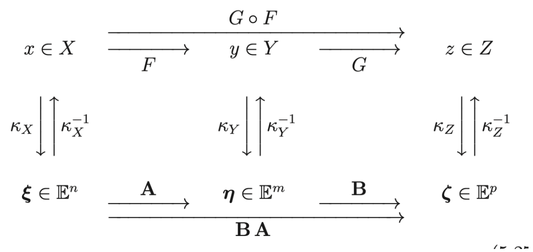
For maps generally:
Definition: image/range $F : X \to Y$. $F(x)$ is the range/image of F: $F(X) = im\ F :\equiv \{F(x) \in Y | x \in X\} \subseteq Y$.
Definition: surjectivity If $F(X) = Y$, then F is on Y: surjective.
Definition: injectivity $F(x) = F(x') \Rightarrow x = x'$: injective (one-to-one).
Definition: bijectivity If surjective and injective: bijective. If F bijective, the inverse map F^{-1} is defined.

definition, matrix representation

Definition: $F : X \to Y, x \mapsto Fx$ (X and Y vector spaces over $\mathbb{E}$) is called linear if $\forall x, \tilde{x} \in X$ and $\forall\gamma, \beta \in \mathbb{E}$: $F(\beta x + \gamma\tilde{x}) = \beta Fx + \gamma F\tilde{x}$:

- $F(x + \tilde{x}) = Fx + F\tilde{x}$
- $F(\gamma x) = \gamma(Fx)$

X is domain, Y is image space/codomain.
If $X = Y$ one has a self-map. If $Y = \mathbb{E}$, F is called linear functional. If X and Y function spaces, F is called linear operator.
Examples are evaluation map, differential operator, multiplication operator.
Arbitrary $F : X \to Y$. $dim\ X = n$ and $dim\ Y = m$. $\{b_1, ..., b_n\}$ basis of X. $\{c_1, ..., c_m\}$ basis of Y.
Definition: Images of basis of X ($Fb_l \in Y$) as linear combination of c_k: $Fb_l = \sum_{k=1}^{m} a_{kl} c_k, l = 1, ..., n$.
$\mathbf{A}_{m\times n} = (a_{kl})$ is called matrix for F relative to the given bases in X and Y.
Notice, l-th column of $\mathbf{A}$: cooridnates of image of l-th basis vector of X relative to chosen basis of Y.
Also, to every $\mathbf{A} \in \mathbb{E}^{m\times n}$ corresponds a unique F for given bases.
IMPORTANT: We have $\eta = \mathbf{A}\zeta$ (neu=$\mathbf{A}$alt).
Definition isomorphism: If $F : X \to Y$ is (eineindeutig), it is an isomorphism. If also $X = Y$: automorphism.
Lemma: 5.1 $F : X \to Y$ isomorphism $\Rightarrow F^{-1} : Y \to X$ also linear and isomorphism.
Definition coordinate mapping: $\kappa_X : X \to \mathbb{E}^n, x \mapsto \zeta$ is bijective and linear, thus an isomorphism. It assigns each $x \in X$, its coordinate vector regarding some basis $\mathcal{B}$.
This commutative diagram helps:



Corollary 5.2: F isomorphism. For fixed bases as $\mathbf{A}$. Then, $\mathbf{A}$ regular and inverse map F^{-1} and $\mathbf{A}^{-1}$.
Theorem 5.3: X, Y, Z vector spaces over $\mathbb{E}$. $F : X \to Y$ and $G : Y \to Z$ linear. Then, $G \circ F : X \to Z$ also linear. If $\mathbf{A}, \mathbf{B}$ map matrices for fixed bases i X, Y, Z for F, G. Then for $G \circ F$: $\mathbf{BA}$.

kernel, image, rank

$F : X \to Y, dim\ X = n, dim\ Y = m$
Definition : kernel of F - $ker\ F$ - is inverse image of $o \in Y$: $ker\ F :\equiv \{x \in X; Fx = o\} \subseteq X$.
Lemma 5.4: $ker\ F$ is subspace of X. $im\ F$ is subspace of Y.
Lemma 5.5: U subspace of $X \Rightarrow FU$ subspace of Y && W subpsace of $im\ F \Rightarrow F^{-1}W$ subspace of W.

- $ker\ \mathbf{A}$ = general solution of the homogenous SLE $\mathbf{Ax} = \mathbf{o}$.
- $im\ \mathbf{A}$ = set of right-hand sides $\mathbf{b}$ for which $\mathbf{Ax} = \mathbf{b}$ has solution.

Theorem 5.6: F is injective if and only if $ker\ F = \{o\}$.
Theorem 5.7 - dimension formula: Assuming $dim\ X < \infty$: $dim\ X - dim\ ker\ F = dim\ im\ F$.

Definition rank: rank of linear map F is dimension of image of F: $rank\ F :\equiv dim\ im\ F$.
Corollary 5.8:

1. $F : X \to Y$ injective $\Leftrightarrow rank\ F = dim\ X$
2. $F : X \to Y$ bijective (isomophism) $\Leftrightarrow rank\ F = dim\ X = dim\ Y$
3. $F : X \to X$ bijective (automorphism) $\Leftrightarrow rank\ F = dim\ X$

Definition: Two vector spaces are called isomorph if an isomorphism $F : X \to Y$ exists.
Theorem 5.9: Two vector spaces with finite dimension are isomorph if and only if they have the same dimension.
Corollary 5.10: $F : X \to Y, G : Y \to Z$ linear maps with $dim\ X, dim\ Y < \infty$.

1. $rank\ FG \leq min\{rank\ F, rank\ G\}$
2. G injective $\Rightarrow rank\ GF = rank\ F$
3. F surjective $\Rightarrow rank\ GF = rank\ G$

matrices as linear maps

$\mathbf{A} = (a_{kl})$, $m \times n$ matrix, n columns: $\mathbf{a}_1, ..., \mathbf{a}_n$.
Definition: space spanned by columns of $\mathbf{A}$ - $\mathcal{R}(\mathbf{A}) :\equiv span\ \{\mathbf{a}_1, ..., \mathbf{a}_n\}$ - is called column space or range of $\mathbf{A}$. The solution space $\mathcal{L}_0$ of the homogenous SLE $\mathbf{Ax} = \mathbf{o}$ is called nullspace of $\mathbf{A}$: $\mathcal{N}(\mathbf{A})$.
Theorem 5.11: If $\mathbf{A}$ is understood as a linear map: $im\ \mathbf{A} = \mathcal{R}(\mathbf{A})$ and $ker\ \mathbf{A} = \mathcal{N}(\mathbf{A})$. $\mathbf{Ax} = \mathbf{b}$ has a solution if and only if $b \in \mathcal{R}(\mathbf{A})$. A solution is unique if and only if $\mathcal{N}(\mathbf{A}) = \{\mathbf{o}\}$.
Theorem 5.12: $dim\ \mathcal{L}_0 \equiv dim\ \mathcal{N}(\mathbf{A}) \equiv dim\ ker\ \mathbf{A} = n - r$.
Definition row space: The space spanned $\mathbb{E}^n$ by row vectors of $m \times n$ is called row space of $\mathbf{A}$.
Theorem 5.13: rank of $\mathbf{A}_{m \times n}$:

- amount of pivot elements in row echelon form
- rank of linear map $\mathbf{A} : \mathbb{E}^n \to \mathbb{E}^m$ as $dim\ im\ \mathbf{A}$
- dimension of column space (column rank)
- dimension of row space (row rank)

Corollary 5.14: $rank\ \mathbf{A}^\top = rank\ \mathbf{A}^H = rank\ \mathbf{A}$.
Theorem 5.15: Column space of $\mathbf{A}_{m \times n}$: $im\ \mathbf{A} \equiv \mathcal{R}(\mathbf{A}) = \mathcal{R}(\tilde{\mathbf{A}}) = span\ \{\mathbf{a}_{n1}, ..., \mathbf{a}_{nr}\}$ with $\mathbf{a}_{nk}$ as pivot columns of $\mathbf{A}$ and $\tilde{\mathbf{A}}$ as $m \times r$ matrix from those.
Theorem 5.16: $\mathbf{A} \in \mathbb{E}^{m \times n}, \mathbf{B} \in \mathbb{E}^{p \times m}$:

- $rank\ \mathbf{BA} \leq min\{rank\ \mathbf{B}, rank\ \mathbf{A}\}$
- $rank\ \mathbf{B} = m (\leq p) \Rightarrow rank\ \mathbf{BA} = rank\ \mathbf{A}$
- $rank\ \mathbf{A} = m (\leq n) \Rightarrow rank\ \mathbf{BA} = rank\ \mathbf{B}$

Corollary 5.17: $\mathbf{A} \in \mathbb{E}^{m \times m}, \mathbf{B} \in \mathbb{E}^{m \times m}$:

- $rank\ \mathbf{BA} \leq min\{rank\ \mathbf{B}, rank\ \mathbf{A}\}$
- $rank\ \mathbf{B} = m \Rightarrow rank\ \mathbf{BA} = rank\ \mathbf{A}$
- $rank\ \mathbf{A} = m \Rightarrow rank\ \mathbf{BA} = rank\ \mathbf{B}$

Theorem 5.18: Equivalent for $\mathbf{A} \in \mathbb{E}^{n \times n}$

- $\mathbf{A}$ is invertible
- $\mathbf{A}$ is regular
- $rank\ \mathbf{A} = n$
- the n column vectors of $\mathbf{A}$ are linearly independent
- the n row vectors of $\mathbf{A}$ are linearly independent
- $im\ \mathbf{A} \equiv \mathcal{R}(\mathbf{A}) = \mathbb{E}^n$
- $ker\ \mathbf{A} \equiv \mathcal{N}(\mathbf{A}) = \{\mathbf{o}\}$
- map $\mathbf{A} : \mathbb{E}^n \to \mathbb{E}^n$ is an automorphism
- $\mathbf{A}$ transformation matrix of coordinate transformation in $\mathbb{E}^n$

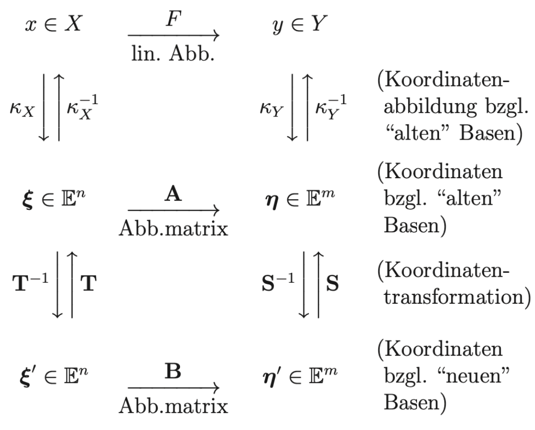
affine spaces, general solution inhomogenous SLE

Definition affine (sub)space: $U \subset V$, $u_0 \in V$. $u_0 + U :\equiv \{u_0 + u | u \in U\}$ is called affine (sub)space.
Definition affine mapping: $F : X \to Y$ linear map and $y_0 \in Y$. $H : X \to y_0 + Y, x \mapsto y_0 + Fx$ is called affine map.
Theorem 5.19: x_0 any solution of $\mathbf{Ax} = \mathbf{b}$. $\mathcal{L}_0$ the general solution of $\mathbf{Ax} = \mathbf{o}$. Then, general solution $\mathcal{L}_\mathbf{b}$ of $\mathbf{Ax} = \mathbf{b}$ is affine subspace $\mathcal{L}_\mathbf{b} = \mathbf{x}_0 + \mathcal{L}_0$.

map matrix for coordiante transforamtion

X and Y vector spaces with dimension n and m.

- $F : X \to Y, x \mapsto y$ - linear map
- $\mathbf{A} : \mathbb{E}^n \to \mathbb{E}^m, \zeta \mapsto \eta$ - some transformation matrix for F
- $\mathbf{T} : \mathbb{E}^n \to \mathbb{E}^n, \zeta' \to \zeta$ - transformation matrix in $\mathbb{E}^n$
- $\mathbf{S} : \mathbb{E}^m \to \mathbb{E}^m, \eta' \mapsto \eta$ - transformation matrix in $\mathbb{E}^m$



For $\mathbf{B}$ of F in new basis in $\mathbb{E}^m$ and $\mathbb{E}^n$: $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AT}$ and $\mathbf{A} = \mathbf{SBT}^{-1}$. With $rank\ F = rank\ \mathbf{A} = rank\ \mathbf{B}$.
Definition similarity: $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are similar if some regular $\mathbf{T}$ exists so that $\mathbf{B} = \mathbf{T}^{-1}\mathbf{AT}$ and $\mathbf{A} = \mathbf{TBT}^{-1}$. $\mathbf{A} \mapsto \mathbf{B} = \mathbf{T}^{-1}\mathbf{AT}$ is called similarity transformation.
Theorem 5.20: $F : X \to Y$ linear map. $dim\ X = n$, $dim\ Y = m$, $rank\ F = r$. Then, transformation matrix $\mathbf{A} = \left(\begin{array}{c|c} \mathbf{I}_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right)$.

vector spaces with scalar product

normed vector spaces

Definition norm: For some vector space V. A norm is a function $\|.\| : V \to \mathbb{R}, x \mapsto \|x\|$ with three characteristics:

1. positiv definit: $\|x\| \geq 0\ \forall x \in V$ & $\|x\| = 0 \Rightarrow x = 0$
2. homogenous in the absolut value: $\|\alpha x\| = |\alpha| \|x\|\ \forall x \in V, \alpha \in \mathbb{E}$
3. triangle inequality: $\|x + y\| \leq \|x\| + \|y\|\ \forall x, y \in V$

V with a norm: normed vector space/normed linear space

vector spaces with scalar product

Definition scalar product: Scalar product in real or complex vector space is a function $\langle ., . \rangle : V \times V \to \mathbb{E}, x, y \mapsto \langle x, y \rangle$ with:

1. linear in second factor: $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle\ \forall x, y, z \in V$ & $\langle x, \alpha y \rangle = \alpha \langle x, y \rangle\ \forall x, y \in V, \alpha \in \mathbb{E}$.
2. hermitian: $\langle x, y \rangle = \overline{\langle y, x \rangle}\ \forall x, y \in V$ (symmetric for real)
3. positiv definit: $\langle x, x \rangle \geq 0\ \forall x \in V$ & $\langle x, x \rangle = 0 \Rightarrow x = 0$

V with scalar product: vector space with scalar/inner product.
Definition: V finite dimension:

- $\mathbb{E} = \mathbb{R}$: Euclidean vector space / orthogonal vector space
- $\mathbb{E} = \mathbb{C}$: unitary vector space

Definition induced norm: Induced norm/length of $x \in V$ with scalar product: $\|x\| :\equiv \sqrt{\langle x, x \rangle}$.
Theorem 6.1 - Cauchy-Bunjakovski-Schwarz inequality: $|\langle x, y \rangle| \leq \|x\| \|y\|$ for all $x, y \in V$. Equality if and only if x, y linearly dependent.
Definition angle: Angle φ $(0 \leq \varphi \leq \pi)$ between $x, y \in V$: $\varphi :\equiv \arccos \frac{\mathrm{Re}\langle x, y \rangle}{\|x\| \|y\|}$. $x, y \in V$ are orthogonal (perpendicular) if and only if $\langle x, y \rangle = 0, x \perp y$. $M, N \subseteq V$ are orthogonal if and only if $\langle x, y \rangle = 0$ for all $x \in M, y \in N$: $M \perp N$.
Theorem 6.2 - pythagorean theorem: $x \perp y \Rightarrow \|x \pm y\|^2 = \|x\|^2 + \|y\|^2$ for all x, y in a vector space with scalar product.

orthonormal bases

Theorem 6.3: Set M of pairwise orthogonal vectors is linearly independent if $\mathbf{o} \notin M$.
Definition: Basis is orthogonal if basis vectors pairwise orthogonal: $\langle b_k, b_l \rangle = 0$ if $k \neq l$. It is orthonormal if additionally length is 1: $\langle b_k, b_k \rangle = 1$ for all k.
Definition Kronecker symbol: $\delta_{kl} :\equiv \left\{\begin{smallmatrix} 0, k \neq l \\ 1, k = l \end{smallmatrix}\right.$. Thus: $\langle b_k, b_l \rangle = \delta_{kl}$.
Theorem 6.4: V n-dimensional vector space with scalar product and orthonormal basis$\{b_1, ..., b_n\}$. For all $x \in V$: $x = \sum_{k=1}^n \langle b_k, x \rangle b_k$. This means: For the coordinates for some orthonormal basis we simply have $\zeta_k = \langle b_k, x \rangle$.
$x = \sum_{k=1}^n \langle b_k, x \rangle b_k = \sum_{k=1}^n b_k (b_k^H x) = (\sum_{k=1}^n b_k b_k^H) x$ Thus: $\mathbf{I}_n = \sum_{k=1}^n b_k b_k^H$.
Theorem 6.5 - Parseval's Theorem: $\zeta_k :\equiv \langle b_k, x \rangle$ and $\eta_k :\equiv \langle b_k, y \rangle$ $(k = 1, ..., n)$. $\langle x, y \rangle = \sum_{k=1}^n \overline{\zeta_k} \eta_k = \zeta^H \eta = \langle \zeta, \eta \rangle$. Thus, the scalar product of two vectors equals the euclidean scalar product of its coordinate vectors. Hence: $\|x\| = \|\zeta\|, \angle(x, y) = \angle(\zeta, \eta)$, $x \perp y \Leftrightarrow \zeta \perp \eta$.
Algorithm 6.1 - Gram-Schmidt process: $\{a_1, a_2, ...\}$ finite or countably finite set of vectors. We compute a same-sized set $\{b_1, b_2, ...\}$:

- $b_1 :\equiv \frac{a_1}{\|a_1\|}$
- $\tilde{b}_k :\equiv a_k - \sum_{j=1}^{k-1} \langle b_j, a_k \rangle b_j$
- $b_k :\equiv \frac{\tilde{b}_k}{\|\tilde{b}_k\|}$

for $k = 2, 3, ...$
Theorem 6.6: The vectors $b_1, b_2, ...$ computed with Gram-Schmidt are normed and pairwise orthogonal. After k septs: $span\{a_1, a_2, ..., a_k\} = span\{b_1, b_2, ..., b_k\}$. If $\{a_1, a_2, ...\}$ is a basis of V $\Rightarrow \{b_1, b_2, ...\}$ is an orthonormal basis of V.
Corollary 6.7: For a vector space with finite or countably finite dimension, an orthonormal basis exists.

orthogonal comlement

Corollary 6.8: In a finite-dimensional (our countably dimensional) vector space with scalar product, every set of orthonormal vectors can be extended to an orthonormal basis.
Definition: V finite-dimensional with scalar product. $U \subset V$. $U^\perp$ (U perp) is orthogonal subspace/orthogonal complement of U. We have $V = U \oplus U', U \perp U^\perp$. Explicitly: $U^\perp :\equiv \{x \in V | x \perp U\}$. V then is the direct sum of orthogonal complements.
Remember $(U^\perp)^\perp = U$ and $dim\ U^\perp + dim\ U = dim\ V$.
Theorem 6.9: $m \times n$ matrix. $\mathbf{A}$ with rank r:

- $\mathcal{N}(\mathbf{A}) = (\mathcal{R}(\mathbf{A}^H))^\perp \subset \mathbb{E}^n$
- $\mathcal{N}(\mathbf{A}^H) = (\mathcal{R}(\mathbf{A}))^\perp \subset \mathbb{E}^m$
- $\mathcal{N}(\mathbf{A}) \oplus \mathcal{R}(\mathbf{A}^H) = \mathbb{E}^n$
- $\mathcal{N}(\mathbf{A}^H) \oplus \mathcal{R}(\mathbf{A}) = \mathbb{E}^m$
- $dim\ \mathcal{R}(\mathbf{A}) = r$
- $dim\ \mathcal{R}(\mathbf{A}^H) = r$
- $dim\ \mathcal{N}(\mathbf{A}) = n - r$
- $dim\ \mathcal{N}(\mathbf{A}^H) = m - r$

Definition: The two paris $\mathcal{N}(\mathbf{A}), \mathcal{R}(\mathbf{A}^H)$ and $\mathcal{N}(\mathbf{A}^H), \mathcal{R}(\mathbf{A})$ are called the four fundamental subspaces of $\mathbf{A}$.

orthogonal/unitary base change

Theorem 6.10: Transformation matrix if change of basis between orthonormal bases is unitary ($\mathbb{E} = \mathbb{C}$) or orthogonal ($\mathbb{E} = \mathbb{R}$). - $\mathbf{I} = \mathbf{T}^H \mathbf{T}$.
Theorem 6.11: Orthogonal/unitary change of basis. Old (ζ) and new (ζ') coordinate vectors linked: $\zeta = \mathbf{T}\zeta'$ and $\zeta' = \mathbf{T}^H \zeta$.
For V some $\mathbb{E}^n$: Basis vector as columns of orthogonal/unitary matrices: $\mathbf{B} = \mathbf{B}' \mathbf{T}^H, \mathbf{B}' = \mathbf{BT}$.

Theorem 6.12: η and η' pair of old/new coordinates: $\langle\zeta',\eta'\rangle = \langle\zeta,\eta\rangle$. Especially: $\|\zeta'\| = \|\zeta\|, \angle(\zeta',\eta') = \angle(\zeta,\eta), \zeta' \perp \eta' \Leftrightarrow \zeta \perp \eta$.

orthogonal/unitary maps

Definition: X, Y unitary/orthogonal vector spaces. $F: X \to Y$ unitary/orthogonal if $\langle Fx, Fy\rangle_Y = \langle x,y\rangle_X$ for any $x, y \in X$.
Theorem 6.13: $F: X \to Y$ orthogonal/unitary.

1. $\|Fx\|_Y = \|x\|_X$ (length preserving/isometric)
2. $x \perp y \Rightarrow Fx \perp Fy$ (angle preserving)
3. $ker\ F = \{o\}$ - F is injective

If $dim\ X = dim\ Y < \infty$. Also:

4. F is isomorphism
5. $\{b_1, ..., b_n\}$ orthonormal basis of X, $\{Fb_1, ..., Fb_n\}$ orthonormal basis of Y.
6. F^{-1} unitary/orthogonal.
7. $\mathbf{A}$ unotariy/orthogonal for orthonormal bases in X, Y

Lemma 6.14: $F: X \to Y, G: Y \to Z$ two unitary/orthogonal isomorphisms of finite dimensional vector spaces with scalar product, so $G \circ F: X \to Z$.
Lemma 6.15: V n-dimensional vector space with scalar product with orthonormal basis, $\kappa_V: V \to \mathbb{E}^n$ is unitary/orthogonal isomorphism.
Lemma 6.16: $\mathbf{A} \in \mathbb{E}^{n\times n}$ is unitary/orthogonal if and only if, $\mathbf{A}: \mathbb{E}^n \to \mathbb{E}^n$ is unitary/orthogonal.

operators and matrices

Definition: X, Y vector spaces with norms $\|.\|_X, \|.\|_Y$. $F: X \to Y$ (linear map/operator) is called bounded if $\gamma_F \geq 0$ with $\|F(x)\|_Y \leq \gamma_F \|x\|_X$ for all $x \in X$. All such maps F between X, Y: $\mathcal{L}(X, Y)$.

ToDo!!! Script pages 148 ff. & Obsidian

least square and QR decomposition

orthogonal projections

Definition: $\mathbf{P}: \mathbb{E}^m \to \mathbb{E}^m$ is called projection or projector of $\mathbf{P}^2 = \mathbf{P}$. $\mathbf{P}$ is called orthogonal projection if $ker\ \mathbf{P} \perp im\ \mathbf{P}$ or $\mathcal{N}(\mathbf{P}) \perp \mathcal{R}(\mathbf{P})$. Otherwise the projection is oblique.
Lemma 7.1: $\mathbf{P}$ projector $\Rightarrow$ $\mathbf{I} - \mathbf{P}$ projector and $im\ (\mathbf{I}-\mathbf{P}) = ker\ \mathbf{P}$ and $ker\ (\mathbf{I}-\mathbf{P}) = im\ \mathbf{P}$
Theorem 7.2: $\mathbf{P}$ porojection. Equivalent:

- $\mathbf{P}$ orthogonal projector
- $\mathbf{I} - \mathbf{P}$ orthogonal projector
- $\mathbf{P}^H = \mathbf{P}$

Lemma 7.3: $\mathbf{A}_{m\times n}, rank\ \mathbf{A} = n (\leq m) \Rightarrow \mathbf{A}^H\mathbf{A}$ is regular
Theorem 7.4: orthogonal projection $\mathbf{P_A}: \mathbb{E}^m \to im\ \mathbf{A} \subseteq \mathbb{E}^m$ on column space $\mathcal{R}(\mathbf{A}) \equiv im\ \mathbf{A}$ of $\mathbf{A}_{m\times n}$ with $rank\ \mathbf{A} = n (\leq m)$: $\mathbf{P_A} :\equiv \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$.
Corollary 7.5: orthogonal projection $\mathbf{P_Q}: \mathbb{E}^m \to im\ \mathbf{Q} \subseteq \mathbb{E}^m$ on column space $\mathcal{R}(\mathbf{Q}) \equiv im\ \mathbf{Q}$ of $\mathbf{Q}_{m\times n} = (\mathbf{q}_1 \ \dots \ \mathbf{q}_n)$ with orthonormal columns: $\mathbf{P_Q} :\equiv \mathbf{Q}\mathbf{Q}^H$. Thus: $\mathbf{P_Q} = \sum_{j=1}^n \mathbf{q}_j\langle\mathbf{q}_j, \mathbf{y}\rangle$.
With the pythagoren theorem, this can be reasoned:
Theorem 7.6: orthogonal projection $\mathbf{P}$. $\|\mathbf{y} - \mathbf{P}\mathbf{y}\|_2 = \min_{\mathbf{z} \in im\ \mathbf{P}} \|\mathbf{y} - \mathbf{z}\|_2$.
Analogous also with different scalar products, because the pythagorean theorem still holds.

least squares

Definition: $\mathbf{Ax} = \mathbf{y}, \mathbf{A}_{m\times n}, m > n$ - overdetermined linear system. A solution only exists if $y \in \mathcal{R}(\mathbf{A})$. If not solution exists, one chooses $\mathbf{x} \in \mathbb{E}^n$ so that the residual/residual vector $\mathbf{r} :\equiv \mathbf{y} - \mathbf{Ax}$ has minimal Euclidean norm (2-norm/length). Such $\mathbf{x}$ is called least square solution of $\mathbf{Ax} = \mathbf{y}$.
Assuming columns of $\mathbf{A}$ to be linearly independent ($ker\ \mathbf{A} = \{o\}$, $\mathbf{A}^H\mathbf{A}$ regular): $\mathbf{x} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{y}$ and $\mathbf{A}^H\mathbf{Ax} = \mathbf{A}^H\mathbf{y}$. Those are called normal equations.
Definition: $(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ is called pseudo-inverse.
Theorem 7.7: $\mathbf{A} \in \mathbb{E}^{m\times n}$, $rank\ \mathbf{A} = n \leq m$, $y \in \mathbb{E}^m$. The overdetermined SLE $\mathbf{Ax} = \mathbf{y}$ has a unique solution $\mathbf{x}$ in the sense of the least square problem: $\|\mathbf{Ax} - \mathbf{y}\|^2 = \min_{\tilde{\mathbf{x}} \in \mathbb{E}^n} \|\mathbf{A}\tilde{\mathbf{x}} - \mathbf{y}\|^2$. $\mathbf{x}$ may be computed by solving the regular system of the normal equations. The residual vector is orthogonal to $\mathcal{R}(\mathbf{A})$.
Lemma 7.8: Let $\mathbf{a}_{n+1} :\equiv \mathbf{y}$. Do Gram-Schmidt on $\mathbf{a}_1, ..., \mathbf{a}_n, \mathbf{a}_{n+1}$. Then: $\tilde{\mathbf{q}}_{n+1} :\equiv \mathbf{y} - \mathbf{Ax} = r \perp \mathcal{R}(\mathbf{A}) = span\{\mathbf{a}_1, ..., \mathbf{a}_n\}$. The system $\mathbf{Ax} = \mathbf{y} - \tilde{\mathbf{q}}_{n+1}$ is then uniquely solvable for $\mathbf{x}$.

QR decomposition

Consider $\mathbf{A}$ with linearly independent columns $\mathbf{a}_1, ..., \mathbf{a}_n$. Remember the Gram-Schmidt process. We may write $\mathbf{a}_1 = \mathbf{q}_1\|\mathbf{a}_1\|$ and $\mathbf{a}_k = \mathbf{q}_k\|\tilde{\mathbf{q}}_k\| + \sum_{j=1}^{k-1} \mathbf{q}_j\langle\mathbf{q}_j, \mathbf{a}_k\rangle$. We define the coefficients $r_{11} = \|\mathbf{a}_1\|$, $r_{jk} = \langle\mathbf{q}_j, \mathbf{a}_k\rangle$ $(j = 1, ..., k-1)$, $r_{kk} :\equiv \|\tilde{\mathbf{q}}_k\|$ for $k = 2, ..., n$. We add $r_{jk} = 0$ $(j = k+1, ..., n)$. We can then write: $\mathbf{a}_k = \mathbf{q}_k r_{kk} + \sum_{j=1}^{k-1} \mathbf{q}_r r_{jk} = \sum_{j=1}^{k} \mathbf{q}_j r_{jk} = \sum_{j=1}^{n} \mathbf{q}_j r_{jk}$.
We then may define $\mathbf{A} :\equiv (\mathbf{a}_1 \ \dots \ \mathbf{a}_n)$ and $\mathbf{Q} :\equiv (\mathbf{q}_1 \dots \mathbf{q}_n)$ and $\mathbf{R} :\equiv \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & r_{nn} \end{pmatrix}$.
Then, we may rewrite the last formula to $\mathbf{A} = \mathbf{QR}$.
Definition QR decomposition: The decomposition above of $\mathbf{A}_{m\times n}$ with rank $n \leq m$ in a $m \times n$ matrix $\mathbf{Q}$ with orthonormal columns and a $n \times n$ upper triangular matrix with positive diagonal elements is called $\mathbf{QR}$ decomposition of $\mathbf{A}$.

We may extend $\mathbf{Q}$ to an orthonormal basis of $\mathbb{E}^n$: $\tilde{\mathbf{Q}} :\equiv (\mathbf{Q}|\mathbf{Q}_\perp) :\equiv (q_1 \ \dots \ q_n | q_{n+1} \ \dots \ q_m)$. We may extend $\mathbf{R}$ with $m - n$ zero columns: $\tilde{\mathbf{R}} :\equiv \binom{\mathbf{R}}{\mathbf{O}}$. Then: $\mathbf{A} = \mathbf{QR} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}$.
Definition: The just introduced 'extended' form is sometimes called $\mathbf{QR}$ decomposition and the earlier mentioned form $\mathbf{QR}$ factorization.
Lemma 7.9: $\mathbf{a}_1, ..., \mathbf{a}_n$ columns of $\mathbf{A}_{\times n}$. Gram-Schmidt leads to $\mathbf{QR}$ decomposition. Adding $\mathbf{y}$ to $\mathbf{A}$ leads to the residual $\mathbf{r} \perp \mathcal{R}(\mathbf{A})$: $\mathbf{r} = \mathbf{y} - \sum_{j=1}^n \mathbf{q}_j\langle\mathbf{q}_j, \mathbf{y}\rangle = \mathbf{y} - \mathbf{Q}\mathbf{Q}^H\mathbf{y}$. For $\mathbf{x}$ as the least square solution, we have $\mathbf{Rx} = \mathbf{Q}^H\mathbf{y}$.

QR decomposition with pivoting

ToDo!!! Obsidian

determinants

permutations

Definition permutation: A permutation of n elements is a unique map of $\{1, ..., n\}$ onto itself. The set of all such permutations is S_n (symmetric group).
This formula is unusable for practical use, because it can only be computed in $\mathcal{O}(n! \cdot n)$. It would take unreasonable amount of time to compute determinants in this way (about 75 years for $n = 20$).
Definition transposition: Permutation with only two elements switched.
Theorem 8.1: There are $n!$ permutations in S_n.
Theorem 8.2: For $n > 1$, every permutation p can be expressed as product of transpositions t_k of neighboring elements: $p = t_\nu \circ t_{\nu-1} \circ ... \circ t_2 \circ t_1$. This is normally not unique. But the number of transpositions is.
Definition sign: permutation p. $sign\ p = \begin{cases} +1, \nu \text{ event} \\ -1, \nu \text{ uneven} \end{cases}$

definition, characteristics

Definition determinant: $\mathbf{A}_{n\times n}$. Determinant: $det\ \mathbf{A} := \sum_{p\in S_n} sign\ p \cdot a_{1,p(1)}a_{2,p(2)} \cdots a_{n,p(n)}$ for all $n!$ permutations.
Theorem 8.3: The determinant is a function $det: \mathbb{E}^{n\times n} \to \mathbb{E}, \mathbf{A} \mapsto det\ \mathbf{A}$ with three characteristics:

1. linearity in each row. For all $\gamma, \gamma' \in \mathbb{E}$ and $l \in \{1, ..., n\}$:
$$det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ \gamma a_{l1} + \gamma' a_{l1} & \dots & \gamma a_{ln} + \gamma' a_{ln} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \gamma det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} + \gamma' det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$
2. switching two rows changes the sign of $det(\mathbf{A})$
3. $det(\mathbf{I}) = 1$

Theorem 8.4: Further characteristics:

4. $\mathbf{A}$ has zero column: $det\ (\mathbf{A}) = 0$
5. $det(\gamma\mathbf{A}) = \gamma^n det(\mathbf{A})$
6. $\mathbf{A}$ has two identical rows: $det(\mathbf{A}) = 0$
7. Adding multiply of one row to another row does not change $det(\mathbf{A})$.
8. $\mathbf{A}$ diagonal matrix: $det(\mathbf{A})$ = product of diagonal elements
9. $\mathbf{A}$ triangular matrix: $det(\mathbf{A})$ = product of diagonal elements

Notice for the Gauss algorithm: Sign of determinant changes when rows switched. Otherwise unchanged.
Theorem 8.5: $\mathbf{A}_{n\times n}$: $det\ \mathbf{A} \neq 0 \Leftrightarrow rank\ \mathbf{A} = n \Leftrightarrow \mathbf{A}$ regular. Applying the Gauss-algorithm to $\mathbf{A}$, ν being the number of row changes: $det\ \mathbf{A} = (-1)^\nu \prod_{k=1}^n r_{kk}$.
Algorithm 8.1: Apply Gauss-algorithm to $\mathbf{A}_{n\times n}$. Then, the formula of theorem 8.5 holds.
Algorithm 8.1 is usually much faster than using the permutation definition. That's because Gauss elimination works in $\mathcal{O}(n^3)$, compared to about $\mathcal{O}(n! \cdot n)$ of the implicit computation of the definition/permutation formula.
Theorem 8.6: The defined determinant is the only function with characteristics (1)-(3).
Theorem 8.7: $\mathbf{A}, \mathbf{B} \in \mathbb{E}^{n\times n}$. $det(\mathbf{AB}) = det(\mathbf{A}) \cdot det(\mathbf{B})$.
Theorem 8.8: $\mathbf{A}_{n\times n}$ regular. $det(\mathbf{A})^{-1} = det(\mathbf{A}^{-1})$.
Theorem 8.9: $det(\mathbf{A}^\top) = det(\mathbf{A})$ and $det(\mathbf{A}^H) = \overline{det(\mathbf{A})}$
Corollary 8.10: Characteristics (1),(2),(4),(6),(7) are also valid for columns (instead of rows).

expansion by rows and columns

Definition: $\mathbf{A}_{n\times n}$. For a_{kl}, we define $(n-1)\times(n-1)$ matrix $\mathbf{A}_{[k,l]}$ by removing row k and colum l from $\mathbf{A}$. Cofactor $\kappa_{kl} :\equiv (-1)^{k+l} det\ \mathbf{A}_{[k,l]}$.
Lemma 8.11: $\mathbf{A}$. Only $a_{kl} \neq 0$ not zero in l column. Then $det\ \mathbf{A} = a_{kl}\kappa_{kl}$.
Theorem 8.12: $\mathbf{A}_{n\times n}$. For all $k, l \in \{1, ..., n\}$: $det\ \mathbf{A} = \sum_{i=1}^n a_{ki}\kappa_{ki}$ (expansion along row k) and $det\ \mathbf{A} = \sum_{i=1}^n a_{il}\kappa_{il}$ (expansion along colum l).

block triangular matrices

Theorem 8.13: For a 2×2 block matrix $det\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{pmatrix} = det\ \mathbf{A} \cdot det\ \mathbf{D}$ or rather $det\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = det\ \mathbf{A} \cdot det\ \mathbf{D}$
Corollary 8.14: The determinant of a block matrix is the product of the determinants of its diagonal blocks.

eigenvalues and eigenvectors

Intuition: We search vectors, which give directions in which some transformation only scales the space but not rotates it. Those are eigenvectors. The scaling factor is the eigenvalue.
V finite dimensional, $F: V \to V, x \mapsto Fx$

eigenvalues/eigenvectors of matrices/linear maps

Definition eigenvalue/eigenvector: $\lambda \in \mathbb{E}$ is eigenvalue of F if an eigenvector $v \in V, v \neq 0$ with $Fv = \lambda v$ exists.
Definition eigenspace: λ eigenvalue. $E_\lambda :\equiv \{v \in V | Fv = \lambda v\}$. Eigenspace is set of eigenvectors with zero vector.
Definition spectrum: Set of all eigenvalues of F. Denoted $\sigma(F)$.

Lemma 9.1: $F : V \to V$ linear map, $\kappa_V : V \to \mathbb{E}^n, x \mapsto \zeta$ coordinate map of V (for some basis), $\mathbf{A} = \kappa_V F \kappa_V^{-1}$ transformation matrix. Then: λ eigenvalue and $\mathbf{x}$ eigenvector of $F \Leftrightarrow \lambda$ eigenvalue and ζ eigenvector of $\mathbf{A}$.
v for λ is not unique, because $Fv = \lambda v \Leftrightarrow F(\alpha v) = \lambda(\alpha v)$. Thus, $dim\ E_\lambda \geq 1$.
If $(F - \lambda I)v = o$ has a non-trivial solution λ is an eigenvalue. E_λ is the general solution.

Lemma 9.2: λ eigenvalue of $F : V \to V$ if and only if $F - \lambda I$ has non-trivial kernel. $E_\lambda = ker\ (F - \lambda I)$. $E_\lambda \neq \{o\}, E_\lambda \subseteq V$.

Definition geometric multiplicity: Geometric multiplicity of λ is dimension of E_λ.

Corollary 9.3: λ eigenvalue of $\mathbf{A} \in \mathbb{E}^{n \times n}$ if $\mathbf{A} - \lambda\mathbf{I}$ singular. $E_\lambda = ker\ (\mathbf{A} - \lambda\mathbf{I}) \neq \{o\}$. Geometric multiplicity of λ: $dim\ E_\lambda = dim\ ker\ (\mathbf{A} - \lambda\mathbf{I}) = n - rank\ (\mathbf{A} - \lambda\mathbf{I})$.

Definition characteristic polynomial and equation: $\mathcal{X}_\mathbf{A}(\lambda) :\equiv det\ (\mathbf{A} - \lambda\mathbf{I})$ is characteristic polynomial of $\mathbf{A} \in \mathbb{E}^{n \times n}$. $\mathcal{X}_\mathbf{A}(\lambda) = 0$ is the characteristic equation.

Definition trace: The sum of the diagonal elements of $\mathbf{A}$: $trace\ \mathbf{A} :\equiv a_{11} + a_{22} + ... + a_{nn}$.

Lemma 9.4: $\mathcal{X}_\mathbf{A}(\lambda) \equiv: det\ (\mathbf{A} - \lambda\mathbf{I}) = (-\lambda)^n + trace\ \mathbf{A}(-\lambda)^{n-1} + ... + \det\ (\mathbf{A})$

Theorem 9.5: $y \in \mathbb{E}$ is eigenvalue of $\mathbf{A} \in \mathbb{E}^{n \times n}$ if and only if λ root of characteristic polynomial/solution of characteristic equation.
According to the fundamental theorem of linear algebra, the characteristic polynomial of degree n has n (usually complex) roots. Becuase then complex eigenvectors exists, we usually consider matrices with real entries also as complex matrices.

Definition algebraic multiplicity: The algeboraic multiplicity of some eigenvalue λ is the multiplicity of λ as root of $\mathcal{X}_\mathbf{A}$.
Notice: algebranic and geometric multiplicity are not necessarily equal!

Algorithm 9.1 - eigenvalue/-vector via $\mathcal{X}_\mathbf{A}$: $\mathbf{A} \in \mathcal{C}^{n \times n}$.

1. Compute $\mathcal{X}_\mathbf{A} :\equiv \det\ (\mathbf{A} - \lambda\mathbf{I})$
2. Compute roots $\lambda_1, ..., \lambda_n$ of $\mathcal{X}_\mathbf{A}$ (with their algebraic multiplicity).
3. For each λ_k: Determine basis of $\ker(\mathbf{A} - \lambda\mathbf{I}) = E_{\lambda_k}$.

Lemma 9.6: $\mathbf{A} \in \mathbb{E}^{n \times n}$ is singular $\Leftrightarrow 0 \in \sigma(\mathbf{A})$.

similarity transformation - spectral decomposition

$F : V \to V, x \mapsto Fx$. $\mathbf{A}, \mathbf{B}$ transformation matrices regarding two different bases. Then, $\mathbf{A}$ and $\mathbf{B}$ are similar. $\mathbf{A} \to \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ is called similarity transformation. We ask: How far $\mathbf{A}$ may be simplified by chosing an appropriate similarity transformation.

Theorem 9.7: Similar matrices have the same characteristic polynomial. Thus, they have the same determinant, trace, eigenvalue. Also the geometric and algebraic multiplicities for some λ is identical for similar matrices.

Lemma 9.8: A transformation matrix for $F : V \to V$ is diagonal if and only if the chosen basis of V consists only of eigenvectors.

Definition: A basis of eigenvectors of F (or $\mathbf{A}$) is an eigenbasis of F (or $\mathbf{A}$).

Theorem 9.9: For $\mathbf{A} \in \mathbb{E}^{n \times n}$ a similar diagonal matrix $\mathbf{\Lambda}$ exists if and only if an eigenbasis exists for $\mathbf{A}$. For $\mathbf{V} :\equiv (\mathbf{v}_1 \quad ... \quad \mathbf{v}_n)$ with $(\mathbf{v}_1, ..., \mathbf{v}_n)$ as the eigenbasis, we have $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}$ and $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$. Accordingly, if some $\mathbf{V}$ and $\mathbf{\Lambda}$ exists, the diagonal elements of $\mathbf{\Lambda}$ are eigenvalues/the columns of $\mathbf{V}$ are eigenvectors of $\mathbf{A}$.

Definition spectral decomposition: $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ (with diagonal $\mathbf{\Lambda}$) is called spectral/eigenvalue decomposition of $\mathbf{A}$. If for some $\mathbf{A}$ such a $\mathbf{\Lambda}$ exists, $\mathbf{A}$ is diagonalizable.

Corollary 9.10: $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$. $\mathbf{V} = (\mathbf{v}_1 \quad ... \quad \mathbf{v}_n)$. $\mathbf{V}^{-1} = \begin{pmatrix}\mathbf{w}_1^\top & ... & \mathbf{w}_n^\top\end{pmatrix}^\top$. Then: $\mathbf{A} = \sum_{k=1}^n \mathbf{v}_k\lambda_k\mathbf{w}_k^\top$. Meanwhile: $\mathbf{A}\mathbf{v}_k = \mathbf{v}_k\lambda$ and $\mathbf{w}_k^\top\mathbf{A} = \lambda\mathbf{w}_k^\top$.

Definition: w is called left eigenvector of $\mathbf{A}$ if $\mathbf{w}^\top\mathbf{A} = \lambda_k\mathbf{w}^\top$.

Theorem 9.11: Eigenvectors for different eigenvalues are linearly independent.

Corollary 9.12: If the n eigenvalues of $F : V \to V$ are distinct ($n = \dim \mathbf{V}$), an eigenbasis exists. The corresponding transformation map is diagonal.

Theorem 9.13: For each eigenvalue, geometric multiplicity $\leq$ algebraic multiplicity.

Theorem 9.14: A matrix is diagonalizable if and only if for each eigenvalue geometric multiplicity = algebraic multiplicity.

symmetric / hermitian matrices

Many eigenvalue problems are self-adjoint - meaning the matrices are real symmetric/hermitian.

Theorem 9.15 - spectral theorem: $\mathbf{A} \in \mathbb{C}^{n \times n}$ hermitian ($\mathbf{A}^H = \mathbf{A}$)

1. All eigenvalues $\lambda_1, ..., \lambda_n$ are real.
2. Eigenvectors for different eigenvalues are pairwise orthogonal in $\mathbb{C}^n$.
3. $\exists$ orthonormal basis of $\mathbb{C}^n$ of eigenvectors $\mathbf{u}_1, ..., \mathbf{u}_n$ of $\mathbf{A}$.
4. $\mathbf{U} :\equiv (\mathbf{u}_1 \quad ... \quad \mathbf{u}_n)$ (unitary). $\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{\Lambda} :\equiv diag\ (\lambda_1, ..., \lambda_n)$

Corollary 9.16: $\mathbf{A} \in \mathbb{R}^{n \times n}$ real ($\mathbf{A}^\top = \mathbf{A}$)

1. All eigenvalues $\lambda_1, ..., \lambda_n$ are real.
2. The real eigenvectors for different eigenvalues are pairwise orthogonal in $\mathbb{R}^n$.
3. $\exists$ orthonormal basis of $\mathbb{R}^n$ of eigenvectors $\mathbf{u}_1, ..., \mathbf{u}_n$ of $\mathbf{A}$.
4. $\mathbf{U} :\equiv (\mathbf{u}_1 \quad ... \quad \mathbf{u}_n)$ (orthogonal). $\mathbf{U}^\top\mathbf{A}\mathbf{U} = \mathbf{\Lambda} :\equiv diag\ (\lambda_1, ..., \lambda_n)$

Definition: $\mathbf{A}$ hermitian.

- positiv definit: $\forall x \in \mathbb{E}^n, x \neq 0$ we have $x^H A x > 0$
- positiv semidefinit: $\forall x \in \mathbb{E}^n$ we have $x^H A x \geq 0$.

Notice, such $\mathbf{A}$ defines a scalar product in $\mathbb{E}^n$: $f : \mathbb{E}^n \times \mathbb{E}^n, (u, v) \mapsto f(u, v) := \mathbf{u}^H\mathbf{A}\mathbf{v}$.

Theorem: $\mathbf{A} \in \mathbb{E}^{n \times n}$ hermitian.

- $\mathbf{A}$ positiv definit $\Leftrightarrow$ all eigenvalues of $\mathbf{A} > 0$
- $\mathbf{A}$ positiv semidefinit $\Leftrightarrow$ all eigenvalues of $\mathbf{A} \geq 0$

Theorem: $\mathbf{A} \in \mathbb{E}^{n \times n}$ is normal $\Leftrightarrow \mathbf{A}$ is diagonalizable by a unitary matrix over $\mathbb{C}$.

Jordan canonical form

Not done.

singular value decomposition

Theorem 11.1: $\mathbf{A} \in \mathbb{C}^{m \times n}$, $rank\ \mathbf{A} = r$. $\exists$ unitary $\mathbf{U} = (\mathbf{u}_1 \quad ... \quad \mathbf{u}_m)$ and unitary $\mathbf{V} = (\mathbf{v}_1 \quad ... \quad \mathbf{v}_n)$ and $\mathbf{\Sigma}_{m \times n} :\equiv \begin{pmatrix}\mathbf{\Sigma}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O}\end{pmatrix}$ with $\mathbf{\Sigma}_r :\equiv diag\ (\sigma_1, ..., \sigma_r)$ with $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > \sigma_{r+1} = ... = \sigma_{min\{m,n\}} = 0$ (positive and ordered). So that: $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \sum_{k=1}^r \mathbf{u}_k\sigma_k\mathbf{v}_k^H$.
Columns of $\mathbf{U}$ orthonormal eigenbasis of $\mathbf{A}\mathbf{A}^H$ ($\mathbf{A}\mathbf{A}^H = \mathbf{U}\mathbf{\Sigma}_m^2\mathbf{U}^H$) ($\mathbf{U}$ diagonalizes $\mathbf{A}\mathbf{A}^H$).
Columns of $\mathbf{V}$ orthonormal eigenbasis of $\mathbf{A}^H\mathbf{A}$ ($\mathbf{A}^H\mathbf{A} = \mathbf{V}\mathbf{\Sigma}_n^2\mathbf{V}^H$) ($\mathbf{V}$ diagonalizes $\mathbf{A}^H\mathbf{A}$). We also have: $\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}$ and $\mathbf{A}^H\mathbf{U} = \mathbf{V}\mathbf{\Sigma}^\top$.
Furthermore:

- $\{\mathbf{u}_1, ..., \mathbf{u}_1\}$ is basis of $im\ \mathbf{A} \equiv \mathcal{R}(\mathbf{A})$
- $\{\mathbf{v}_1, ..., \mathbf{v}_r\}$ is basis of $im\ \mathbf{A}^H \equiv \mathcal{R}(\mathbf{A}^H)$
- $\{\mathbf{u}_{r+1}, ..., \mathbf{u}_m\}$ is basis of $ker\ \mathbf{A}^H \equiv \mathcal{N}(\mathbf{A}^H)$
- $\{\mathbf{v}_{r+1}, ..., \mathbf{v}_n\}$ is basis of $ker\ \mathbf{A} \equiv \mathcal{N}(\mathbf{A})$

Those are all orthonormal bases. If $\mathbf{A}$ is real, $\mathbf{U}, \mathbf{V}$ may be chosen as real orthogonal matrices.

Definition singular value decomposition: The matrix factorization $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = \sum_{k=1}^r \mathbf{u}_k\sigma_k\mathbf{v}_k^H$ (from above) is called singular value decomposition (SVD) of $\mathbf{A}$. $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > \sigma_{r+1} = ... = \sigma_{min\{m,n\}} = 0$ are called singular values of $\mathbf{A}$. $\mathbf{u}_1, ..., \mathbf{u}_m$ are called left singular vectors. $\mathbf{v}_1, ..., \mathbf{v}_n$ are called right singular vectors.

Corollary 11.2: $\mathbf{V}_r$ matrix with first r columns of $\mathbf{V}$. $\mathbf{U}_r$ matrix with first r columns of $\mathbf{U}$. $\mathbf{\Sigma}_r$ leading $r \times r$ matrix of $\mathbf{\Sigma}$. Compact SVD: $\mathbf{A} = \mathbf{U}_r\mathbf{\Sigma}\mathbf{V}_r^H = \sum_{k=1}^r \mathbf{u}_k\sigma_k\mathbf{v}_k^H$. $\mathbf{U}_{m \times r}$ and $\mathbf{V}_{n \times r}$ with orthonormal columns. Diagonal elements of Σ_r positive.

derivation

$A^H A$ is diagonalizable (according to the spetral theorem) and all eigenvalues of $A^H A$ are ≥ 0, because its positiv semidefinit.
spectral decomposition of $A^H A$ exists because of spectral theorem: $A^H A V = V\Lambda$, $\Lambda = diag(\lambda_1, ..., \lambda_n)$, V is unitary, $\lambda_j \geq 0$ $(j = 1, ..., n)$. We may sort the eigenvalues: $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_r > \lambda_{r+1} = ... = \lambda_n = 0$ (notice, $r = rank\ (A^H A) = rank\ A$). And define $\sigma_j := \sqrt{\lambda_j}$. This is possible, because $\lambda_j \geq 0$.
Then: $A^H A (v_1 \quad ... \quad v_r \quad v_{r+1} \quad ... \quad v_n) = (v_1 \quad ... \quad v_r \quad v_{r+1} \quad ... \quad v_n) \begin{pmatrix}\sigma_1^2 & & & & & \\ & ... & & & & \\ & & \sigma_2^2 & & & \\ & & & 0 & & \\ & & & & ... & \\ & & & & & 0\end{pmatrix}$ We may simplify with $A^H A V_r = V_r \Sigma_r$: $A^H A (v_1 \quad ... \quad v_r) = (v_1 \quad ... \quad v_r) \begin{pmatrix}\sigma_1^2 & & \\ & ... & \\ & & \sigma_2^2\end{pmatrix}$ For Σ_r we have $\Sigma_r^{-1} = \begin{pmatrix}\frac{1}{\sigma_1^2} & & \\ & ... & \\ & & \frac{1}{\sigma_2^2}\end{pmatrix}$ and we know that $V_r^H V_r = I_{r \times r}$, because the columns of V_r are orthonormal. We get: $A^H A V_r = V_r \Sigma_r^2$ $V_r^H A^H A V_r = \Sigma_r^2$ $(\Sigma_r^{-1} V_r^H a^H)(A V_r \Sigma_r^{-1}) = I_{r \times r}$ (multiplying with Σ_r^{-1} left and right) $U_r^H U_r = I_{r \times r}$ with $U_r := A V_r \Sigma_r^{-1} \Rightarrow A V_r = U_r \Sigma_r$. Here, U_r is $m \times r$, Σ_r is $r \times r$, and V_r is $n \times r/V_r^{-1}$ is $r \times n$. This is the reduced SVD.
We can extend $U_r \in \mathbb{E}^{m \times r}$ to a unitary matrix $U \in \mathbb{E}^{m \times m}$: $U := (U_r | U_r^\top)$ with $U^H U = U U^H = I_{m \times m}$.
Then we have $AV = U\Sigma \Rightarrow A = U\Sigma V^H$.

optional/additional stuff

Givens rotations

script, page 63

Householder matrices/reflections

script, page 64

permutation matrices

script, page 64