Cheat Sheet for MATH461

Here is the stuff you really need to remember for the exams.

1 Linear systems $Ax = b$

Problem:

We consider a linear system of m equations for n unknowns $x_1, \dots, x_n$:
For a given matrix $A \in \mathbb{R}^{m \times n}$ and a vector $b \in \mathbb{R}^m$ we want to find all vectors $x \in \mathbb{R}^n$ such that $Ax = b$.

Algorithms:

Gaussian elimination: This gives the decomposition $\boxed{\begin{bmatrix} \text{row } p_1 \text{ of } A \\ \vdots \\ \text{row } p_m \text{ of } A \end{bmatrix} = LU}$ where

- the **row echelon form** $U \in \mathbb{R}^{m \times n}$ has r nonzero rows and r nonzero pivots in columns $q_1, \dots, q_r$. The variables $x_{q_1}, \dots, x_{q_r}$ are called **basic variables**, the remaining variables are called **free variables**.
- the matrix $L \in \mathbb{R}^{m \times m}$ is lower triangular: the diagonal elements are 1, below the diagonal are the multipliers used in the elimination
- the permuation vector p contains the numbers $1, \dots, m$ in a scrambled order

For given L, U, p

- **find basis for** range$A$**:** Use columns $q_1, \dots, q_r$ of A
- **find basis for** null$A$**:** Set one of the free variables to 1 and the others to zero. Then use $Ux = \vec{0}$ and back substitution to find the basic variables $x_{q_r}, \dots, x_{q_1}$. This gives vectors $v^{(1)}, \dots, v^{(n-r)}$.
- **find basis for** range$A^\top$**:** Use rows $1, \dots, r$ of U
- **find basis for** null$A^\top$**:** For $j = 1, \dots, m-r$: Solve $L^\top u = e^{(r+j)}$ by forward substititition. Then let $\begin{cases} w_{p_1} := u_1 \\ \quad\vdots \\ w_{p_m} := u_m \end{cases}$. This gives $m - r$ basis vectors $w^{(1)}, \dots, w^{(m-r)}$. Here $e^{(k)} := [\overset{1}{0}, \dots, \overset{k-1}{0}, \overset{k}{1}, \overset{k+1}{0}, \dots, \overset{m}{0}]^\top$.

For given L, U, p and right hand side vector $b \in \mathbb{R}^m$ **find general solution of linear system** $Ax = b$**:**

- solve $Ly = \begin{bmatrix} b_{p_1} \\ \vdots \\ b_{p_m} \end{bmatrix}$ by forward substitution
- find particular solution x_{part} of $Ux = y$: Set all free variables to 0. Then use back substitution to find the basic variables $x_{q_r}, \dots, x_{q_1}$.
- Find a basis $v^{(1)}, \dots, v^{(n-r)}$ for null$A$ using U (see above)
- The general solution is $x = x_{\text{part}} + c_1 v^{(1)} + \dots + c_{n-r} v^{(n-r)}$ where $c_1, \dots, c_{n-r} \in \mathbb{R}$ are arbitrary

Subspace $V = \operatorname{span}\left\{v^{(1)},\dots,v^{(k)}\right\}$ with given vectors $v^{(j)} \in \mathbb{R}^n$

- **find basis for V:**
  - Let A have the rows $v^{(1)\top},\dots,v^{(k)\top}$, then $V = \operatorname{range} A^\top$
  - Find row echelon form U (we don't need L, p)
  - use rows $1,\dots,r$ of U
- **find basis for orthogonal complement $V^\perp$:**
  - since $V^\perp = \operatorname{null} A$: use row echelon form U to find basis for $\operatorname{null} A$

Important facts to remember:

- $\operatorname{range} A$ and $\operatorname{range} A^\top$ have both dimension r
- $\operatorname{null} A$ and $\operatorname{range} A^\top$ are orthogonal complements in $\mathbb{R}^n$, $\dim \operatorname{null} A = n - r$
- $\operatorname{null} A^\top$and $\operatorname{range} A$ are orthogonal complements in $\mathbb{R}^m$, $\dim \operatorname{null} A^\top = m - r$
  Hence: $Ax = b$ has a solution $\iff$ b is orthogonal on basis vectors of $\operatorname{null} A^\top$ (we have $m - r$ solvability conditions)

Case of square matrix $A \in \mathbb{R}^{n\times n}$

- If $r = \operatorname{rank} A < n$ we say that A is **singular**. If $r = n$ we say A is **nonsingular**
- A is singular $\iff$ $\det A = 0$ $\iff$ rows of A are linearly dependent $\iff$ columns of A are linearly dependent
- If A is singular: linear system $Ax = b$ has either no solution or infinitely many solutions
- If A is nonsingular: linear system $Ax = b$ has a unique solution $x \in \mathbb{R}^n$.
  Let $u^{(j)}$ denote the solution vector of $Au^{(j)} = e^{(j)}$, then $A^{-1} = \left[u^{(1)},\dots,u^{(n)}\right] \in \mathbb{R}^{n\times n}$ is called **inverse matrix**.
  Then the solution of $Ax = b$ is given by $x = A^{-1}b$.

2 Least squares problem $\|Ac - b\| = \min$

Problem:

We have a matrix $A \in \mathbb{R}^{m\times n}$ with $m \geq n$ and linearly independent columns. Since the linear system $Ac = b$ does in general not have a solution, we want to find $c \in \mathbb{R}^n$ such that $\| \underbrace{Ac - b}_{\text{residual}\, r} \|$ is as small as possible.

Application:

"Curve fitting": We have measured experimental data $\dfrac{t_1 \quad \cdots \quad t_m}{y_1 \quad \cdots \quad y_m}$ and want to fit this with a curve

$$y = c_1 g_1(t) + \cdots + c_n g_n(t)$$

Here $g_1(t),\dots,g_n(t)$ are given functions and $c_1,\dots,c_n$ are unknown parameters which we want to find.

We want that the residuals $r_j = c_1 g_1(t_j) + \cdots + c_n g_n(t_j) - y_j$ are "small":

Least squares fit: find $c_1,\dots,c_n$ such that $r_1^2 + \cdots + r_m^2$ is minimal.

With $A = \begin{bmatrix} g_1(t_1) & \cdots & g_m(t_1) \\ \vdots & & \vdots \\ g_1(t_m) & \cdots & g_n(t_m) \end{bmatrix}$ we have $r = Ac - y$. Hence we want to find $c \in \mathbb{R}^n$ such that $\|Ac - y\|$ is minimal.

Algorithms:

- Method 1: **Normal equations:** Solve the $n \times n$ linear system $\boxed{(A^\top A)c = A^\top b}$.
- Method 2: **Orthogonalization**
  - Find orthogonal basis $p^{(1)}, \dots, p^{(n)}$ for span$\left\{a^{(1)}, \dots, a^{(n)}\right\}$ using **Gram-Schmidt process**:

$$p^{(1)} := a^{(1)}, \qquad p^{(2)} := a^{(2)} - \underbrace{\frac{a^{(2)} \cdot p^{(1)}}{p^{(1)} \cdot p^{(1)}}}_{s_{12}} p^{(1)}, \qquad p^{(3)} := a^{(3)} - \underbrace{\frac{a^{(3)} \cdot p^{(1)}}{p^{(1)} \cdot p^{(1)}}}_{s_{13}} p^{(1)} - \underbrace{\frac{a^{(3)} \cdot p^{(2)}}{p^{(2)} \cdot p^{(2)}}}_{s_{23}} p^{(2)}, \quad \dots$$

    yielding a decomposition $\boxed{A = PS}$ where

$$\underset{m \times n}{P = \left[p^{(1)}, \dots, p^{(n)}\right]} \text{ has orthogonal columns, } S = \underset{n \times n}{\begin{bmatrix} 1 & s_{12} & \cdots & s_{1n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & s_{n-1,n} \\ 0 & \cdots & 0 & 1 \end{bmatrix}} \text{ is upper triangular.}$$

  - Let $d_j := \dfrac{p^{(j)} \cdot b}{p^{(j)} \cdot p^{(j)}}$ for $j = 1, \dots, n$
  - solve $Sc = d$ by back substitution

3 The determinant $\det A$

For a square matrix $A \in \mathbb{R}^{n \times n}$

- $\det A = 0 \iff$ matrix is singular (i.e. columns are linearly dependent)
- expansion formula: Use **row 1** or the **column 1** to write $\det A$ in terms of $(n-1) \times (n-1)$ determinants:

$$\begin{aligned} \det A &= a_{11} \det A_{[11]} - a_{12} \det A_{[12]} + \dots + (-1)^{n+1} a_{1n} \det A_{[1n]} \\ &= a_{11} \det A_{[11]} - a_{21} \det A_{[21]} + \dots + (-1)^{n+1} a_{n1} \det A_{[n1]} \end{aligned}$$

  Here $A_{[jk]}$ denotes the matrix where we remove row j and column k.
  This also works with the **row** j or **column** j of the matrix, but with the factor $(-1)^{j-1}$.

- $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc, \qquad \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} = a \cdot \det \begin{bmatrix} e & f \\ h & k \end{bmatrix} - b \cdot \det \begin{bmatrix} d & f \\ g & k \end{bmatrix} + c \cdot \begin{bmatrix} d & e \\ g & h \end{bmatrix}$

4 Eigenvalue problem $Av = \lambda v$

4.1 Finding eigenvalues and eigenvectors

Problem:

A square matrix $A \in \mathbb{R}^{n \times n}$ describes a mapping $\mathbb{R}^n \to \mathbb{R}^n$. If we choose a new basis $v^{(1)}, \dots, v^{(n)}$ of $\mathbb{R}^n$ this mapping is represented by the matrix $B = V^{-1}AV$ where $V = \left[v^{(1)}, \dots, v^{(n)}\right]$. We want to find a new basis such that the matrix B becomes "as simple as possible": if possible, we would like B to be a diagonal matrix.

Hence want to find a vectors $v \neq \vec{0}$ and a numbers λ such that $Av = \lambda v$.

Algorithm:

How to find eigenvalues:

- find characteristic polynomial $p(\lambda) = \det \underbrace{\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn}-\lambda \end{bmatrix}}_{A-\lambda I} = (-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots +$ $c_1\lambda + c_0 = p(\lambda)$ by evaluating the determinant
- solve $p(\lambda) = 0$ to find the eigenvalues $\lambda_1, \ldots, \lambda_n$.
  For $n = 2$ you have to solve a quadratic equation. For $n = 3$ try to guess one eigenvalue λ_1, then you can write $p(\lambda) = (\lambda - \lambda_1)q(\lambda)$ with a quadratic polynomial $q(\lambda)$. I will give you hints how to find the eigenvalues if $n > 2$.

How to find eigenvectors: For each eigenvalue $\lambda_1, \ldots, \lambda_n$ solve the linear system

$$\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \cdots & a_{n,n-1} & a_{nn}-\lambda \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

i.e., **we have to find a basis for** $\text{null}(A - \lambda I)$**.** Since $\det(A - \lambda I) = 0$ we have $r := \text{rank}(A - \lambda I) < n$, and we have to **find** $n - r$ **linearly independent vectors**. We call $\text{null}(A - \lambda I)$ the **eigenspace for the eigenvalue** λ.

By taking the eigenvectors for all eigenvalues we obtain vectors $v^{(1)}, \ldots, v^{(m)}$. These vectors are linearly independent, hence

- either $m = n$: then the matrix is called **diagonizable**, i.e., the vectors $v^{(1)}, \ldots, v^{(n)}$ form a basis for all of $\mathbb{R}^n$
- or $m < n$: then the matrix is called **not diagonizable**

Note: eigenvalues and eigenvectors may be complex.

4.2 Case of diagonizable matrices

If we can find n **linearly independent eigenvectors** $v^{(1)}, \ldots, v^{(n)}$, then $V = \left[v^{(1)}, \ldots, v^{(n)}\right]$ is nonsingular. Then by changing to the new basis $v^{(1)}, \ldots, v^{(n)}$ the matrix A becomes the diagonal matrix $B = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$: (Note that blank elements in the matrix are supposed to be zero)

$$A \underbrace{\left[v^{(1)}, \ldots, v^{(n)}\right]}_{V} = \underbrace{\left[v^{(1)}, \ldots, v^{(n)}\right]}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}_{B}, \qquad A = VBV^{-1}, \qquad V^{-1}AV = B$$

4.3 Case of non-diagonizable matrices and the Jordan normal form

If all eigenvalues are different the matrix is always diagonizable. **If there are multiple eigenvalues** it can happen that there are **fewer eigenvectors than the multiplicity of the eigenvalue**.

Assume, e.g., that $\lambda = 2$ is a triple eigenvalue, but $\dim \text{null}(A - 2I) = 1$, i.e., there is only one eigenvector $v^{(1)}$. Then we can find find two **generalized eigenvectors** $v^{(1,1)}, v^{(1,2)}$ satisfying

$$(A - \lambda I)v^{(1)} = 0, \qquad (A - \lambda I)v^{(1,1)} = v^{(1)}, \qquad (A - \lambda I)v^{(1,2)} = v^{(1,1)}$$

We call $v^{(1)}, v^{(1,1)}, v^{(1,2)}$ a **Jordan chain** of length 3. Note that we have

$$A\left[v^{(1)}, v^{(1,1)}, v^{(1,2)}\right] = \left[v^{(1)}, v^{(1,1)}, v^{(1,2)}\right] \begin{bmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}$$

where instead of a diagonal matrix $\begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{bmatrix}$ we now have a so-called **Jordan box** $\begin{bmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}$.

Example: assume that for $A \in \mathbb{R}^{6\times 6}$ we have $\lambda_1 = \lambda_2 = \lambda_3 = 2$, $\lambda_4 = \lambda_5 = 3$, $\lambda_6 = 5$. If there is only one eigenvector $v^{(1)}$ for $\lambda = 2$ we can find a Jordan chain $v^{(1)}, v^{(1,1)}, v^{(1,2)}$ of length 3. If there is only one eigenvector $v^{(4)}$ for $\lambda = 3$ we can find a Jordan chain $v^{(4)}, v^{(4,1)}$ of length 2. Then we have $V = \left[v^{(1)}, v^{(1,1)}, v^{(1,2)}, v^{(4)}, v^{(4,1)}, v^{(6)}\right]$ and $AV = VB$ with

$$B = \begin{bmatrix} 2 & 1 & & & & \\ & 2 & 1 & & & \\ & & 2 & & & \\ & & & 3 & 1 & \\ & & & & 3 & \\ & & & & & 5 \end{bmatrix}$$

Main result (Jordan normal form): For any matrix $A \in \mathbb{R}^{n\times n}$ we can find a basis of Jordan chains for $\mathbb{R}^n$. If V is the matrix which has these Jordan chains as columns, then we have $AV = VB$ where the matrix B has Jordan boxes along the diagonal.

4.4 Symmetric matrices and quadratic forms

A quadratic form is a sum of terms with $x_i x_j$. We can write it as

$$q(x) = \sum_{\substack{i=1\dots n\\ j=1\dots n}} a_{ij} x_i x_j = x^\top A x$$

with a **symmetric matrix** $A \in \mathbb{R}^{n\times n}$. E.g., consider for $n = 3$ the quadratic form

$$q(x) = x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_1x_3 - 2x_2x_3 = [x_1, x_2, x_3] \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Change of basis: For linearly independent vectors $v^{(1)}, \dots, v^{(n)}$ we can write $x = y_1 v^{(1)} + \dots + y_n v^{(n)} = Vy$ with the matrix $V = \left[v^{(1)}, \dots, v^{(n)}\right]$. Then the quadratic form becomes

$$q(x) = x^\top A x = y^\top \underbrace{V^\top A V}_{B} y = y^\top B y$$

with $B = V^\top A V$.

Main result for a symmetric matrix $A \in \mathbb{R}^{n\times n}$: We can find an **orthonormal basis** $v^{(1)}, \dots, v^{(n)}$ **of eigenvectors** (i.e., $V^\top V = I$ and $V^{-1} = V^\top$) such that $V^\top A V = B = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ with **real eigenvalues** $\lambda_1, \dots, \lambda_n$, hence the quadratic form becomes with the new variables

$$q(x) = x^\top A x = y^\top B y = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2$$

- find characteristic polynomial $p(\lambda) = \det(A - \lambda I)$
- solve $p(\lambda) = 0$, this gives eigenvalues $\lambda_1, \ldots, \lambda_n$ which are real
- for each eigenvalue λ: find a basis for $\text{null}(A - \lambda I)$
  this gives n linearly independent eigenvectors $u^{(1)}, \ldots, u^{(n)}$ (i.e., the matrix A is diagonalizable)
- for multiple eigenvalues: find an orthogonal basis by using the Gram-Schmidt process
  normalize the basis vectors: $v^{(j)} := u^{(j)} / \left\| u^{(j)} \right\|$.

Example: For $A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$ we obtain the eigenvalues $\lambda_1 = -1$, $\lambda_2 = 2$, $\lambda_3 = 2$.

For $\lambda = -1$ we have $A - \lambda I = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$ which has rank 2, hence we get one eigenvector $v^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. For $\lambda = 2$ we have $A - \lambda I = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix}$ which has rank 1, hence we get two eigenvectors $v^{(2)} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, $v^{(3)} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.
Note that $v^{(2)}$and $v^{(3)}$ are not orthogonal on each other. Therefore we can use Gram-Schmidt to replace $v^{(3)}$ with

$$w^{(3)} = v^{(3)} - \frac{v^{(3)} \cdot v^{(2)}}{v^{(2)} \cdot v^{(2)}} v^{(2)} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

Now the three vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$ are orthogonal on each other. We now divide each vector by its norm to obain an orthonormal basis: $\frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \sqrt{\frac{2}{3}} \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$. Hence we obtain

$$V = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \sqrt{\frac{2}{3}} \end{bmatrix}, \qquad B = V^\top A V = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \qquad q(x) = y^\top B y = -y_1^2 + 2y_2^2 + 2y_3^2$$

Positive definite matrices, positive semidefinite matrices

We say a symmetric matrix $A \in \mathbb{R}^{n \times n}$ is **positive definite** if $q(x) = x^\top A x > 0$ for all $x \neq \vec{0}$.

We say a symmetric matrix $A \in \mathbb{R}^{n \times n}$ is **positive semidefinite** if $q(x) = x^\top A x \geq 0$ for all $x \neq \vec{0}$.

Result: A is positive definite $\iff$ all eigenvalues of A are positive
A is positive semidefinite $\iff$ all eigenvalues of A are ≥ 0.

Example: The matrix $A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}$ is not positive definite since it has a negative eigenvalue $\lambda_1 = -1$.