Mathematics 1: First midterm Cheat-Sheet (Theory)

Linear Algebra

Basics

$N(A) = \{\vec{x} \in \mathbb{R}^n; A\vec{x} = \vec{0}\}$,
$C(A) = \mathscr{L}\{A^{(1)}, ..., A^{(n)}\} = \{A\vec{x} : \vec{x} \in \mathbb{R}^n\}$, $N(A^T) = C(A)^T$ and $N(A)^\perp = C(A^T)$.
Eigenvalues - $\Delta A(x) = det(A - xI)$.
Eigenvectors - $A\vec{x} = \lambda\vec{x} \leftrightarrow (A - \lambda I)\vec{x} = 0$, non zero solution $det(A - \lambda I) = 0$ or $rank(A - \lambda I)$ not full and $\vec{x} = N(A - \lambda I)$.
Computing the inverse of a matrix:
$\begin{bmatrix} A & \vert & I \end{bmatrix} \to Gauss \to \begin{bmatrix} I & \vert & A^{-1} \end{bmatrix}$.
$Q \in \mathbb{R}^{n\times m}$ is *orthogonal* $\leftrightarrow Q = [\vec{q}_1, ..., \vec{q}_n]$,
$\vec{q}_i \cdot \vec{q}_j = 0 \leftrightarrow \vec{q}_i^T \vec{q}_j = 0$ (pairwise orthogonal), $||\vec{q}_i|| = 1$.
$Q^T Q = I \leftrightarrow Q^{-1} = Q^T$.
$A \in \mathbb{R}^{2\times 2}$...$det(A) = a_{11}a_{22} - a_{12}a_{21}$.
$A \in \mathbb{R}^{3\times 3}$...$det(A) = a_{11}det(A_{x_1}) - a_{12}det(A_{x_2}) + a_{13}det(A_{x_3})$.
A invertible $\leftrightarrow det(A) \neq 0$, $det(A) = \lambda_1 \cdot ... \cdot \lambda_n$ and $\lambda_i \neq 0$.
A symmetric diagonalized: $A = PDP^{-1} = QDQ^T$.
A symmetric, if eigenvalues are real, eigenvectors are orthogonal, full set of independent eigenvectors.
Full column rank $\leftrightarrow$ columns are lin. independent.

Trace

$tr(A) = \sum_{i=1}^n a_{i,i}$, properties:
$tr(\alpha A) = \alpha tr(A)$, $tr(A + B) = tr(A) + tr(B)$, $tr(A^T) = tr(A)$, $tr(AB) = tr(BA)$, $tr(PAP^{-1}) = tr(A)$ (P invertible), $tr(ABC) = tr(CAB) = tr(BCA)$ (order remains).

Rank

$rank(A) = rk(A)$ is *the number of pivots in reduced row echelon form = number of linearlly independednt rows = dim. of the linear span of rows of A = number of lin. independent columns = dim of the lin. span of columns of A =* $dimC(A)$ $= n - dimN(A) =$ *size of the largest invertible square sumbatrix of A.*

Matrix similarity

$A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ are similiar if exists invertible matrix $P \in \mathbb{R}^{n\times n}$ such that $A = PBP^{-1}$. Similar matrices have the same *trace*, *determinant*, *characteristic polynomial*, *eigenvalues* and *rank*. Matrix is diagonalizable if it is similar to some diagonal matrix $A = PDP^{-1}$. The diagonal values of matrix D are eigenvalues of matrix A and the columns of matrix P are eigenvectors of matrix A. $A, B \in \mathbb{R}^{n\times n}$ are orthogonally similar if $A = QBQ^{-1} = QBQ^T$ where $Q \in \mathbb{R}^{n\times n}$ is orthogonal matrix.

Schur's Theorem

$A \in \mathbb{R}^{n\times n}$ has $\mathbb{R}$ eigenvalues $\lambda_1, ..., \lambda_n$, exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$, that $Q^T AQ$ is upper (or lower) triangular $n \times n$ matrix with diagonal entries λ_i. Matrix A is of the form QDQ^T where D is a diagonal matrix with eigenvalues of A on the diagonal and Q is an orthogonal matrix. If matrix A has eigenvalues $\lambda_1, ..., \lambda_n$ then $tr(A) = \lambda_1 + ... + \lambda_n$ and $det(A) = \lambda_1 ... \lambda_n$.

Frobenius norm

Scalar (inner) product $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$ is $\langle A, B\rangle = tr(A^T B)$ with properties:
$\langle A, B\rangle = \langle B, A\rangle$, $\langle \alpha A + \beta B, C\rangle = \alpha\langle A, C\rangle + \beta\langle B, C\rangle$, for all $\alpha, \beta \in \mathbb{R}$, $\langle A, A\rangle \geq 0$, $\langle A, A\rangle = 0 \leftrightarrow A = 0$, For matrices $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{m\times k}$ and $C \in \mathbb{R}^{k\times n}$ we have $\langle A, BC\rangle = \langle B^T A, C\rangle = \langle AC^T, B\rangle$.
$A = [a_{i,j}] \in \mathbb{R}^{m\times n} \to ||A||_F = ||A|| = \sqrt{\langle A, A\rangle} = \sqrt{tr(A^T A)} = ||vec(A)||$. $\sigma_1, ..., \sigma_k$ are singular values of A, $||A||_F = \sum_{i=1}^{rkA} \sigma_i^2 = tr(A^T A)$, $A^T A \in \mathbb{R}^{n\times n}$, $\lambda_i = \sigma_i^2$.

Kronocker Product

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix} \in \mathbb{R}^{mp\times nq}$$

Properties:
1. $0 \otimes A = A \otimes 0 = 0$
2. $\alpha \otimes A = A \otimes \alpha = \alpha A$, $\forall\, \alpha \in \mathbb{R}$
3. $(\alpha A) \otimes B = A \otimes (\alpha B) = \alpha(A \otimes B)$
4. $(A + B) \otimes C = A \otimes C + B \otimes C$ and $A \otimes (B + C) = A \otimes B + A \otimes C$
5. $(A \otimes B)^T = A^T \otimes B^T$
6. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$
7. $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$
8. $||A \otimes B||_F = ||A||_F ||B||_F$
9. If $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$, then $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$
10. If $A \in \mathbb{R}^{n\times n}$ has eigenvalues $\lambda_1, \dots \lambda_m$ and B has eigenvalues $\mu_1, \dots, \mu_n$ then the set of eigenvalues of $A \otimes B$ is equal to $\{\lambda_i \mu_j\}$
11. If $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$, then $tr(A \otimes B) = tr(A)tr(B)$
12. If $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$, then $det(A \otimes B) = (\det A)^m (\det B)^n$
13. $rank(A \otimes B) = rank(A)\, rank(B)$
14. If $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$ and $C \in \mathbb{R}^{p\times r}$, then $vec(ABC) = (C^T \otimes A)vec(B)$

PSD matrices

Matrix quadratic form

Quadratic form of $A \in \mathbb{R}^{n\times n}$:

$$\vec{x}^T A\vec{x} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j.$$

$A = QDQ^T \to x^T Ax = u^T Du$ where $u = Q^T x$ and $u^T Du = \lambda_1 u_1^2 + ... + \lambda_n u_n^2$.
A symmetric matrix $A \in \mathbb{R}^{n\times n}$ is:
Positive semidefinite (PSD) if $x^T Ax \geq 0 \forall x \in \mathbb{R}^n$ or $\leftrightarrow$ all eigenvalues of A are non-negative.
Positive definite (PD) if $x^T Ax > 0\forall$ nonzero $x \in \mathbb{R}^n$ or $\leftrightarrow$ all eigenvalues of A are positive ($det(A) > 0$).
Negative semidefinite (NSD) if $x^T Ax \leq 0 \forall x \in \mathbb{R}^n$ or $\leftrightarrow$ all eigenvalues of A are non-positive.
Negative definite (ND) if $x^T Ax < 0\forall$ nonzero $x \in \mathbb{R}^n$ or $\leftrightarrow$ all eigenvalues of A are negative.
Indefinite if $x^T Ax > 0$ for some $x \in \mathbb{R}^n$ and $y^T Ay < 0$ for some $y \in \mathbb{R}^n$ or $\leftrightarrow$ A has positive and negative eigenvalues.

Sylvester

A symmetric matrix A is PD if and only if the determinant of each leading principal submatrix is positive and PSD when its non-negative. A symmetric matrix A is ND if and only if the determinant of the $k \times k$ leading principal submatrix is positive if k is even and negative if k is odd.

Cholesky decomposition

QR decomposition: Q is orthogonal matrix of B and R is upper triangular matrix of coefficients.
$A = BB^T = (QR)^T QR = R^T Q^T QR = R^T R = LL^T$.
For invertible, (symmetric) and PSD matrix $A \in \mathbb{R}^{n\times n}$ we have *Decomposition algorithm*:
Write $A_1 := A = \begin{bmatrix} a_{11} & \vec{b}^T \\ \vec{b} & B \end{bmatrix}$, define $L_1 := \begin{bmatrix} \sqrt(a_{11}) & \vec{0}^T \\ \frac{1}{\sqrt{a_{11}}}\vec{b} & I_{n-1} \end{bmatrix}$.
$A_1 = L_1 \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & B - \frac{1}{a_{11}}\vec{b}\vec{b}^T \end{bmatrix} L_1^T$.
Repeat this on $A_2 := B - \frac{1}{a_{11}}\vec{b}\vec{b}^T \in \mathbb{R}^{(n-1)\times(n-1)}$.
If $L_1, L_2, ...L_3$ are the matrices obtained in this way then:
$L = L_1 \cdot \begin{bmatrix} 1 & \vec{0}^T \\ \vec{0} & L_2 \end{bmatrix} \cdot ... \cdot \begin{bmatrix} I_{n-1} & \vec{0}^T \\ \vec{0} & L_n \end{bmatrix}$.
If one of these steps fails then the matrix A is not PSD.

Vector Spaces

Vector space

V is a set of vectors $v \in V$ with two inner operations: addition - $u, v \in V \Rightarrow u + v \in V$ and scalar multiplication - $v \in V, \alpha \in \mathbb{R} \Rightarrow \alpha v = \alpha \cdot v \in V$.
There exists a zero vector 0 and $v + 0 = v$ and for each $v \in V$ exists an inverse vector $-v$, such that $v + (-v) = 0$.
$1 \cdot v = v$, $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$, $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$, $\alpha \cdot (u + v) = \alpha \cdot v + \alpha \cdot u$, for all $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$.
Zero vector 0 is unique, $0v = 0$ and $\alpha 0 = 0$.
Linear combination of vectors: vector of the form $\alpha_1 v_1 + ... + \alpha_n v_n$.

Vector subspaces

Subset U of a VS V is a vector subspace if its *closed under linear combinations* $\alpha u + \beta v \in U$.
Linear span $\mathscr{L}\{v_1, ..., v_n\}$ is the set of all linear combinations. It is the smallest vector subspace containing vectors $v_1, ..., v_n$.

Basis of a vector space

Vectors $v_1, ..., v_n$ are linearly dependent if $\exists v_k$, written as a linear combination
$v_k = \alpha_1 v_1 + ... + \alpha_{k-1} v_{k-1} + \alpha_{k+1} v_{k+1} + ... + \alpha_n v_n$ and *linearly independent* if it doesn't exist. Vectors $v_1, ..., v_n$ are linearly independent if the only linear combination equal to 0 is $\alpha_1 v_1 + ... + \alpha_n v_n = 0 \leftrightarrow \alpha_1 = ... = \alpha_n = 0$.

The set of vectors $\mathscr{B} = \{v_1, ..., v_n\} \subseteq V$ is a *basis* of V if $v_1, ..., v_n$ are linearly independent and $\mathscr{L} = \{v_1, ..., v_n\}$ span V. The number of elements in any basis of vector space V is by $dimV$ ($dim\mathbb{R}^n = n; dim\mathbb{R}^{n\times m} = nm; dim\mathbb{R}_n[x] = n+1$).

Linear transformations

Transformation $\tau : V \to U$ is linear if $\tau(u+v) = \tau(u) + \tau(v)$ and $\tau(\alpha v) = \alpha\tau(v) \leftrightarrow \tau(\alpha v + \beta u) = \alpha\tau(v) + \beta\tau(u)$ holds. $\tau(0) = 0$.

Operations with linear transformations

sum - $\tau + \phi : V \to U$ as $(\tau + \phi)(v) = \tau(v) + \phi(v)$,
multiple - $\gamma\tau : V \to U$ as $(\omega\tau)(v) = \omega\tau(v)$,
composition - $\theta \circ \tau : V \to W$ as $(\theta \circ \tau)(v) = \theta(\tau(v))$.

Matrix corresponding to lin. transf.

Images of vectors $v \in V$ from $\mathscr{B}$ in $\mathscr{C}$ we write $\tau(b_j) = \alpha_{1j}c_1 + ... + \alpha_{nj}c_n$.

$$A_{\tau,\mathscr{B},\mathscr{C}} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & ... & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & ... & \alpha_{2m} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & ... & \alpha_{nm} \end{bmatrix}$$

is *a matrix corresponding to the linear transformation τ from basis $\mathscr{B}$ to basis $\mathscr{C}$*. Columns $= \tau(u_i)$ and rows $= c_j$.
This matrix of a linear transformation has the properties:
$A_{\tau+\phi,\mathscr{B},\mathscr{C}} = A_{\tau,\mathscr{B},\mathscr{C}} + A_{\phi,\mathscr{B},\mathscr{C}}$, $A_{\alpha\tau,\mathscr{B},\mathscr{C}} = \alpha A_{\tau,\mathscr{B},\mathscr{C}}$,
$A_{\theta\circ\tau,\mathscr{B},\mathscr{D}} = A_{\theta,\mathscr{C},\mathscr{D}} \cdot A_{\tau,\mathscr{B},\mathscr{C}}$, $A_{\phi^{-1},\mathscr{B},\mathscr{C}} = (A_{\phi,\mathscr{B},\mathscr{C}})^{-1}$.

Eigenvalues of lin. transf.

All matrices corresponding to linear transformation τ have the same eigenvalues as τ. $\tau(v) = \lambda v = \lambda\mathscr{I}(v)$ where $\mathscr{I}$ is the identity transf. $\mathscr{I} : V \to V$ (or $(\tau - \lambda\mathscr{I})(v) = 0$) $\to$ kernel of $\tau - \lambda\mathscr{I}$. Linear transformation is diagonalizable $\leftrightarrow$ there exists a basis $\mathscr{B}$ of V consisting of eigenvectors of τ.

Kernel and Image

Kernel ($ker(\tau) = ker\tau$) is set of all vectors $v \in V$ $\tau(v) = 0$.
Image is set $im(\tau) = im\tau = \{\tau(v) : v \in V\} \subseteq U$. $ker\tau$ is a vector subspace of V and $im\tau$ is a vector subspace of U.
Linear transformation is injective if and only if $ker\tau = 0$ or $\tau(u) = \tau(v) \Rightarrow u = v$.
Linear transformation is surjective if and only if $im\tau = U$.
$dim(im\tau) = rank(A)$, $dim(ker\tau) + dim(im\tau) = dim(V)$,
$dim(ker\tau) = dim(N(A_\tau))$, $dim(im\tau) = dim(C(A_\tau))$,
$ker\tau \leftrightarrow N(A)$ in basis $\mathscr{B}$ and $im\tau \leftrightarrow C(A)$ in basis $\mathscr{C}$.