Linear Algebra Cheat Sheet

HKPFS Math PhD Interview Prep

1 1. System of Linear Equations

1.1 Matrix Equation Form

System: $a_{11}x_1 + \cdots + a_{1m}x_m = b_1$, ..., $a_{n1}x_1 + \cdots + a_{nm}x_m = b_n$
Matrix form: $Ax = b$ where A is $n \times m$ matrix
Augmented matrix: $(A|b)$

1.2 Solution Methods

Matrix inversion (square invertible): $x = A^{-1}b$
Cramer's rule: For invertible $n \times n$ matrix A:

$$x_i = \frac{\det(A_i)}{\det(A)}$$

where A_i is A with i-th column replaced by b

1.3 Elementary Row Operations

Type I: Interchange two rows
Type II: Multiply row by nonzero scalar
Type III: Add scalar multiple of one row to another
Elementary matrix: Apply operation to I_n
EA = result of applying operation E to A
All elementary matrices are invertible

1.4 Reduced Row Echelon Form (RREF)

Leading entry: First nonzero entry in row
Leading one: Leading entry equals 1
RREF conditions:
1. Leading one is only nonzero in its column
2. All-zero rows at bottom
3. Leading ones shift right as rows descend

Gaussian elimination: Transform to RREF via elementary operations
Leading variable: Corresponds to leading entry
Free variable: Not a leading variable

1.5 Consistency

Consistent: Has solution (solution set nonempty)
Inconsistent: No solution
$Ax = b$ inconsistent $\Leftrightarrow$ RREF of $(A|b)$ has row $[0 \cdots 0|c]$ with $c \neq 0$

1.6 Homogeneous Systems

$Ax = 0$ always has trivial solution $x = 0$
If $m > n$ (more variables than equations), infinite solutions exist
General solution structure: If x_p satisfies $Ax_p = b$, then

$$\text{Solution set} = \{x_p + y : Ay = 0\}$$

1.7 Invertibility and Solutions

For $n \times n$ matrix A:
$Ax = 0$ has nontrivial solution $\Leftrightarrow$ A not invertible
A invertible $\Leftrightarrow$ RREF of A is I_n

2 2. Determinants

2.1 Definition

2×2: $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$
$n \times n$ (cofactor expansion): Fix row i or column j:

$$\det(A) = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} \det(\tilde{A}_{ik})$$

where $\tilde{A}_{ij}$ is (i,j)-minor (delete row i, column j)

2.2 Properties

$\det(I_n) = 1$, $\det(0) = 0$
Transpose: $\det(A^T) = \det(A)$
Switching rows/columns: Changes sign
Two identical rows/columns: $\det(A) = 0$
Scalar multiplication: $\det(cA) = c^n \det(A)$
Multiplicative: $\det(AB) = \det(A)\det(B)$
Row operation: Adding scalar multiple of row to another doesn't change determinant
Invertibility: A invertible $\Leftrightarrow \det(A) \neq 0$
If invertible: $\det(A^{-1}) = \frac{1}{\det(A)}$

2.3 Adjugate and Inverse

Adjugate: $\text{adj}(A) = ((-1)^{i+j} \det(\tilde{A}_{ij}))^T$
Inverse formula: $A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$

2.4 Finding Inverse via Row Operations

Form $[A|I_n]$, row reduce to $[I_n|A^{-1}]$
Every invertible matrix is product of elementary matrices

3 3. Vector Spaces

3.1 Definition

Vector space V over $\mathbb{R}$ with operations $+$ and scalar multiplication satisfying:
1. $x + y = y + x$ (commutative)
2. $(x + y) + z = x + (y + z)$ (associative)
3. $\exists 0 : x + 0 = x$ (zero element)
4. $\forall x, \exists y : x + y = 0$ (additive inverse)
5. $1x = x$
6. $a(bx) = (ab)x$
7. $(a + b)x = ax + bx$
8. $a(x + y) = ax + ay$

3.2 Examples

$\mathbb{R}^n$ with standard operations
$M_{n \times m}$ (all $n \times m$ matrices)
$P_n(\mathbb{R})$ (polynomials degree $\leq n$)
Functions $\mathbb{R} \to \mathbb{R}$ with pointwise operations

3.3 Vector Subspaces

Nonempty $S \subseteq V$ is subspace if:
1. $v_1, v_2 \in S \Rightarrow v_1 + v_2 \in S$ (closed under addition)
2. $v \in S, c \in \mathbb{R} \Rightarrow cv \in S$ (closed under scalar mult)

Examples: Solution set of $Ax = 0$; $\{0\}$; V itself
In $\mathbb{R}^2$: Only $\{0\}$, lines through origin, $\mathbb{R}^2$
In $\mathbb{R}^3$: Only $\{0\}$, lines through origin, planes through origin, $\mathbb{R}^3$

3.4 Linear Combinations

$v = a_1u_1 + \cdots + a_nu_n$ for scalars a_i and vectors u_i
Span: $\text{span}(S) = \{$all linear combinations of vectors in $S\}$
$\text{span}(S)$ is a subspace (smallest subspace containing S)
Column space: $\text{col}(A) = \text{span}\{\text{columns of } A\} = \{Ax : x \in \mathbb{R}^m\}$

3.5 Linear Independence

$S = \{v_1, \ldots, v_r\}$ is **linearly independent** if

$$a_1v_1 + \cdots + a_rv_r = 0 \Rightarrow a_1 = \cdots = a_r = 0$$

Otherwise **linearly dependent**
Test: Solve $a_1v_1 + \cdots + a_rv_r = 0$ (system of linear equations)

3.6 Basis and Dimension

Basis: Linearly independent set that spans V
Every vector in V has unique representation as linear combination of basis vectors
Standard basis for $\mathbb{R}^n$: $\{e_1, \ldots, e_n\}$ where e_i has 1 in position i
Dimension: $\dim(V)$ = number of vectors in any basis
All bases of V have same number of vectors
If $\dim(V) = n$ and S has n linearly independent vectors, then S is basis
If $W \subseteq V$ subspace, $\dim(W) \leq \dim(V)$

4 4. Linear Transformations

4.1 Definition

$T : V \to W$ is **linear transformation** if:
1. $T(x + y) = T(x) + T(y)$
2. $T(cx) = cT(x)$

Consequences: $T(0) = 0$;
$T(a_1x_1 + \cdots + a_rx_r) = a_1T(x_1) + \cdots + a_rT(x_r)$

4.2 Examples

$T(x) = Ax$ for matrix A (most important!)
Differentiation on $C^1(\mathbb{R})$
Integration on continuous functions
Zero transformation $T_0(v) = 0$
Identity $\text{Id}_V(v) = v$

4.3 Construction

Given basis $\{v_1, \ldots, v_n\}$ of V and any $w_1, \ldots, w_n \in W$:
$\exists!$ linear $T : V \to W$ with $T(v_i) = w_i$

4.4 Null Space and Range

Null space: $N(T) = \{v \in V : T(v) = 0\}$ (kernel)
Range: $R(T) = \{T(v) : v \in V\}$ (image)
Both are subspaces
Computing range: If $\{v_1, \ldots, v_m\}$ basis for V:

$$R(T) = \text{span}\{T(v_1), \ldots, T(v_m)\}$$

Nullity: $\text{nullity}(T) = \dim(N(T))$
Rank: $\text{rank}(T) = \dim(R(T))$

4.5 Dimension Formula

$$\text{nullity}(T) + \text{rank}(T) = \dim(V)$$

For matrix A ($n \times m$): - $\text{rank}(A)$ = number of leading ones in RREF - $\text{nullity}(A)$ = number of free variables - $\text{rank}(A) + \text{nullity}(A) = m$

4.6 Injectivity and Surjectivity

Injective: $T(x) = T(y) \Rightarrow x = y$
T injective $\Leftrightarrow N(T) = \{0\} \Leftrightarrow \text{rank}(T) = \dim(V)$
If $\dim(V) > \dim(W)$, then T not injective
Surjective: $R(T) = W$
Isomorphism: Bijective (injective and surjective) linear map
T isomorphism $\Rightarrow \dim(V) = \dim(W)$
Inverse: If T invertible, T^{-1} also linear

5 5. Matrix Representations

5.1 Coordinate Vectors

Ordered basis $\beta = \{v_1, \ldots, v_m\}$ for V
For $v = a_1v_1 + \cdots + a_mv_m$:

$$[v]_\beta = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

5.2 Matrix of Linear Transformation

$T: V \to W$, $\beta = \{v_1, \dots, v_m\}$ basis for V, $\gamma = \{w_1, \dots, w_n\}$ basis for W

$[T]_\gamma^\beta$ is $n \times m$ matrix with j-th column $= [T(v_j)]_\gamma$

Key property: $[T(v)]_\gamma = [T]_\gamma^\beta [v]_\beta$

Standard matrix: For $T: \mathbb{R}^m \to \mathbb{R}^n$, use standard bases

$\text{rank}(T) = \text{rank}([T]_\gamma^\beta)$

5.3 Operations on Transformations

Addition: $(T + T')(v) = T(v) + T'(v)$

$[T + T']_\gamma^\beta = [T]_\gamma^\beta + [T']_\gamma^\beta$

Composition: $(T' \circ T)(v) = T'(T(v))$

$[T' \circ T]_\alpha^\beta = [T']_\alpha^\gamma [T]_\gamma^\beta$

Inverse: $[T^{-1}]_\beta^\gamma = ([T]_\gamma^\beta)^{-1}$

T invertible $\Leftrightarrow$ $[T]_\gamma^\beta$ invertible

5.4 Change of Basis

β, β' two bases for V

Change of coordinate matrix: $Q = [\text{Id}_V]_{\beta'}^\beta$

Q is invertible; $[v]_\beta = Q[v]_{\beta'}$

Similarity: Matrices A, B are similar if $\exists$ invertible Q: $B = Q^{-1}AQ$

For $T: V \to V$ with bases β, β':

$$[T]_{\beta'} = Q^{-1}[T]_\beta Q$$

where $Q = [\text{Id}_V]_{\beta'}^\beta$

Similar matrices have same determinant and rank

6 6. Inner Products

6.1 Definition

Inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ satisfying:

1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
2. $\langle cx, y \rangle = c\langle x, y \rangle$
3. $\langle x, y \rangle = \langle y, x \rangle$ (symmetric)
4. $\langle x, x \rangle > 0$ if $x \neq 0$ (positive definite)

Standard inner product on $\mathbb{R}^n$:

$$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n = x^T y$$

For functions on $[a, b]$: $\langle f, g \rangle = \int_a^b f(x)g(x)dx$

For matrices: $\langle A, B \rangle = \text{tr}(A^T B)$

Weighted inner product: $\langle x, y \rangle = a_1 x_1 y_1 + \dots + a_n x_n y_n$ where $a_i > 0$

Key property: $\langle Ax, y \rangle = \langle x, A^T y \rangle$

6.2 Norm and Distance

Norm (length): $\|x\| = \sqrt{\langle x, x \rangle}$

$\|cx\| = |c|\|x\|$; $\|x\| = 0 \Leftrightarrow x = 0$

Cauchy-Schwarz: $|\langle x, y \rangle| \leq \|x\|\|y\|$

Triangle inequality: $\|x + y\| \leq \|x\| + \|y\|$

6.3 Orthogonality

Orthogonal: $x \perp y$ if $\langle x, y \rangle = 0$

Unit vector: $\|v\| = 1$

Normalize: $\frac{v}{\|v\|}$ is unit vector

Orthogonal set: Pairwise orthogonal; always linearly independent (if nonzero)

Orthonormal set: Orthogonal and all unit vectors

6.4 Orthonormal Basis

Standard basis $\{e_1, \dots, e_n\}$ is orthonormal for $\mathbb{R}^n$

If $\{v_1, \dots, v_r\}$ orthonormal basis and $y = a_1 v_1 + \dots + a_r v_r$:

$$a_i = \langle v_i, y \rangle$$

If orthogonal (not normalized): $a_i = \frac{\langle v_i, y \rangle}{\|v_i\|^2}$

6.5 Gram-Schmidt Process

Given basis $\{w_1, \dots, w_n\}$, construct orthogonal basis $\{v_1, \dots, v_n\}$:

$v_1 = w_1$

$v_i = w_i - \sum_{j=1}^{i-1} \frac{\langle w_i, v_j \rangle}{\|v_j\|^2} v_j$ for $i \geq 2$

Then normalize each v_i to get orthonormal basis: $\left\{ \frac{v_1}{\|v_1\|}, \dots, \frac{v_n}{\|v_n\|} \right\}$

Every inner product space has orthonormal basis

6.6 Orthogonal Complement

For subset $S \subseteq V$:

$$S^\perp = \{x \in V : \langle x, y \rangle = 0 \text{ for all } y \in S\}$$

$S^\perp$ is subspace; $S^\perp = \text{span}(S)^\perp$

Dimension formula: $\dim(W) + \dim(W^\perp) = \dim(V)$

$(W^\perp)^\perp = W$ for subspace W

6.7 Orthogonal Projection

For subspace W with orthonormal basis $\{v_1, \dots, v_r\}$:

Every $v \in V$ uniquely: $v = x + y$ where $x \in W$, $y \in W^\perp$

Projection onto W:

$$\text{proj}_W(v) = \langle v, v_1 \rangle v_1 + \dots + \langle v, v_r \rangle v_r$$

$\text{proj}_W(v)$ minimizes $\|v - w\|$ over all $w \in W$

Finding orthonormal basis for $W^\perp$:

1. Start with orthonormal basis $\{v_1, \dots, v_r\}$ for W
2. Extend to basis $\{v_1, \dots, v_r, w_{r+1}, \dots, w_n\}$ for V
3. Apply Gram-Schmidt to get $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$
4. Then $\{v_{r+1}, \dots, v_n\}$ is orthonormal basis for $W^\perp$

7 7. Least Squares Approximation

7.1 Problem Setup

Given data points $(t_1, y_1), \dots, (t_k, y_k)$

Find line $y = ct + d$ that best approximates data

Equivalently: solve $Ax \approx y$ where

$$A = \begin{pmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_k & 1 \end{pmatrix}, \quad x = \begin{pmatrix} c \\ d \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix}$$

7.2 Least Squares Solution

Find x_0 that minimizes $\|Ax - y\|$ (distance from y to $\text{col}(A)$)

Solution: x_0 satisfies the **normal equation**:

$$A^T A x_0 = A^T y$$

Interpretation: Ax_0 is orthogonal projection of y onto $\text{col}(A)$

$Ax_0 - y \in (\text{col}(A))^\perp = N(A^T)$

7.3 Finding Least Squares Solution

Method 1 (Normal equations):

1. Compute $A^T A$ (always square, symmetric)
2. Compute $A^T y$
3. Solve $(A^T A)x_0 = A^T y$
4. If $A^T A$ invertible: $x_0 = (A^T A)^{-1} A^T y$

Method 2 (Orthogonal projection):

1. Find orthonormal basis $\{v_1, \dots, v_m\}$ for $\text{col}(A)$ (Gram-Schmidt)
2. Compute $\text{proj}_{\text{col}(A)}(y) = \sum_{i=1}^m \langle y, v_i \rangle v_i$
3. Solve $Ax_0 = \text{proj}_{\text{col}(A)}(y)$

7.4 Properties

$A^T A$ is invertible $\Leftrightarrow$ columns of A linearly independent

$\text{rank}(A^T A) = \text{rank}(A)$ (always!)

$\text{nullity}(A^T A) = \text{nullity}(A)$

Least squares solution always exists; unique if columns of A linearly independent

7.5 General Polynomial Fitting

For polynomial $y = a_0 + a_1 t + \dots + a_m t^m$:

$$A = \begin{pmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^m \\ 1 & t_2 & t_2^2 & \cdots & t_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_k & t_k^2 & \cdots & t_k^m \end{pmatrix}, \quad x = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{pmatrix}$$

Then solve normal equation $(A^T A)x = A^T y$

8 8. Eigenvalues & Eigenvectors

8.1 Definitions

For $n \times n$ matrix A:

Eigenvector: Nonzero v such that $Av = \lambda v$ for some scalar λ

Eigenvalue: Scalar λ such that $Av = \lambda v$ for some nonzero v

Eigenspace: $E_\lambda = \{v : Av = \lambda v\} = N(A - \lambda I)$

8.2 Finding Eigenvalues

Characteristic polynomial: $p_A(t) = \det(A - tI)$

Degree n polynomial with leading coefficient $(-1)^n$

λ is eigenvalue $\Leftrightarrow$ $\det(A - \lambda I) = 0$

At most n eigenvalues

8.3 Finding Eigenvectors

For eigenvalue λ: solve $(A - \lambda I)v = 0$

Solution space is E_λ (always contains nonzero vectors)

8.4 Multiplicities

Algebraic multiplicity: Largest k such that $(t - \lambda)^k | p_A(t)$

Geometric multiplicity: $\dim(E_\lambda) = \text{nullity}(A - \lambda I)$

Always: $1 \leq$ geom mult $\leq$ alg mult

8.5 Diagonalization

A is **diagonalizable** if $\exists$ invertible Q: $Q^{-1}AQ = D$ diagonal

Criteria: A diagonalizable $\Leftrightarrow$ $\exists$ basis of $\mathbb{R}^n$ of eigenvectors

$\Leftrightarrow$ For each eigenvalue: geom mult = alg mult

If A has n distinct eigenvalues, A is diagonalizable

How to diagonalize: Find eigenvectors $v_1, \dots, v_n$ (basis)

Set $Q = [v_1 \cdots v_n]$

Then $Q^{-1}AQ = \text{diag}(\lambda_1, \dots, \lambda_n)$ where $Av_i = \lambda_i v_i$

Eigenvectors with distinct eigenvalues are linearly independent

8.6 Applications

Powers: $A = QDQ^{-1} \Rightarrow A^n = QD^nQ^{-1}$
$D^n = \text{diag}(\lambda_1^n, \ldots, \lambda_n^n)$
Matrix exponential: $e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$
If $A = QDQ^{-1}$: $e^A = Qe^DQ^{-1}$ where $e^D = \text{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})$

9 9. Quick Reference

9.1 Matrix Properties

$(AB)^T = B^TA^T$
$(AB)^{-1} = B^{-1}A^{-1}$
$(A^T)^{-1} = (A^{-1})^T$
$\text{rank}(A^TA) = \text{rank}(A)$
$\text{nullity}(A^TA) = \text{nullity}(A)$

9.2 Dimension Results

$\dim(\mathbb{R}^n) = n$
$\dim(M_{n\times m}) = nm$
$\dim(P_n(\mathbb{R})) = n + 1$

9.3 Invertibility Equivalences

For $n \times n$ matrix A, TFAE:
- A invertible - $\det(A) \neq 0$ - RREF of A is I_n - $\text{rank}(A) = n$ - Columns of A linearly independent - Columns of A span $\mathbb{R}^n$ - $Ax = 0$ only has trivial solution - $Ax = b$ has unique solution for any b - A is product of elementary matrices

9.4 Common Mistakes to Avoid

$\det(A + B) \neq \det(A) + \det(B)$ in general

Eigenvectors must be nonzero by definition

Similar matrices have same eigenvalues but not necessarily same eigenvectors

Change of basis matrix depends on order

In least squares: A^TA may not be invertible

Gram-Schmidt must maintain order of vectors