CARD #1

$\nabla(x^T x) = ?$

Hint: SDD

GRADIENT FORMULAS

BASIC

Front

ANSWER #1

$\nabla(x^T x) = 2x$

SDD: Self Dot Doubles

Why?

$x^T x = \Sigma x_i^2$
$d/dx_i = 2x_i$
$\therefore \nabla = 2x$

Use: Finding ∇f when $f(x) = ||x||_2^2$

⚠ Don't forget the 2!

Back

CARD #2

$\nabla(a^T x) = ?$

Hint: LKC

GRADIENT FORMULAS

BASIC

Front

ANSWER #2

$\nabla(a^T x) = a$

LKC: Linear Keeps Constant

Why?

$a^T x = \Sigma a_i x_i$ (linear)
$d/dx_i = a_i$
$\therefore \nabla = a$

Use: Linear functions, constant gradient

⚠ Gradient is a, not x!

Back

CARD #3

$\nabla(x^T Ax) = ?$
(general A)

Hint: QAT

GRADIENT FORMULAS

INTERMEDIATE

Front

ANSWER #3

$\nabla(x^T Ax) = (A + A^T)x$

QAT: Quadratic Adds Transpose

Why?

Expand: $x_i A_{ij} x_j$
d/dx_k gives $A_{kj} x_j + x_i A_{ik}$
$= Ax + A^T x$

Use: Quadratic forms, any matrix A

⚠ If A symmetric: $\nabla = 2Ax$

Back

CARD #4

$\nabla(x^T Ax) = ?$
(A symmetric)

Hint: QAT special

GRADIENT FORMULAS

INTERMEDIATE

Front

ANSWER #4

$\nabla(x^T Ax) = 2Ax$

Symmetric: $A = A^T$
so $(A+A^T)x = 2Ax$

Why?

When $A = A^T$:
$(A + A^T)x = 2Ax$

Use: $A = A^T$ (symmetric matrix)

⚠ Check if A is symmetric first!

Back

CARD #5

$\nabla(||Ax||_2^2) = ?$

Hint: NSDA

GRADIENT FORMULAS

ADVANCED

Front

ANSWER #5

$\nabla(||Ax||_2^2) = 2A^T Ax$

NSDA: Norm Squared Doubles A^T A

Why?

$||Ax||^2 = (Ax)^T(Ax) = x^T A^T Ax$
$A^T A$ is symmetric
$\therefore \nabla = 2A^T Ax$

Use: Least squares, residuals

⚠ It's A^T A, not AA^T!

Back

CARD #6

$||x||_1 = ?$

Hint: LAA (Taxi)

VECTOR NORMS

BASIC

Front

ANSWER #6

$||x||_1 = \Sigma|x_i|$

LAA: L_1 Adds Absolutes

Why?

Manhattan/Taxicab norm
Sum of absolute values

Use: Sparse solutions, L1 regularization

⚠ Don't forget absolute value!

Back

CARD #7

$||x||_2 = ?$

Hint: SSR (Bird)

VECTOR NORMS

BASIC

Front

ANSWER #7

$||x||_2 = \sqrt{(\Sigma x_i^2)} = \sqrt{(x^T x)}$

SSR: Square Sum Root

Why?

Euclidean norm
Straight-line distance

Use: Standard distance metric

⚠ Don't forget square root!

Back

CARD #8

$||x||_\infty = ?$

Hint: MM (Mountain)

VECTOR NORMS

BASIC

Front

ANSWER #8

$||x||_\infty = \max|x_i|$

MM: Maximum/Max

Why?

Infinity norm
Largest component

Use: Worst-case analysis

⚠ It's the MAX, not sum!

Back

CARD #9

Relationship:
$||x||_\infty \; ? \; ||x||_2 \; ? \; ||x||_1$

Hint: Mountain ≤ Bird ≤ Taxi

NORM INEQUALITY

BASIC

Front

ANSWER #9

$||x||_\infty \leq ||x||_2 \leq ||x||_1$

Max ≤ Euclidean ≤ Sum

Why?

Max is smallest measure
Sum is largest

Use: Comparing norms

⚠ Inequality direction!

Back

CARD #10

$(AB)^T = ?$

Hint: RROS

MATRIX TRANSPOSE

BASIC

Front

ANSWER #10

$(AB)^T = B^T A^T$

RROS: Reverse Roll Order Switches

Why?

Order reverses in transpose

Use: Simplifying matrix expressions

⚠ NOT A^T B^T!

Back

CARD #11

Is A^T A always positive definite?

Hint: PSD vs PD

MATRIX PROPERTY

INTERMEDIATE

Front

ANSWER #11

A^T A is always PSD (positive semidefinite)

$x^T(A^T A)x = ||Ax||^2 \geq 0$

Why?

Equals squared norm
Always ≥ 0

Use: Least squares, Gram matrices

⚠ PSD not PD (allows zero)

Back

CARD #12

What is SPD?

Hint: Two properties

POSITIVE DEFINITE

INTERMEDIATE

Front

ANSWER #12

SPD = Symmetric + PD
A = A^T AND x^T Ax > 0

SPD: Symmetric Positive Definite

Why?

Both properties needed

Use: Inner products, convexity

⚠ Need BOTH conditions!

Back

CARD #13

PD Test Method 1 (eigenvalues)

Hint: All positive

PD TESTS

INTERMEDIATE

Front

ANSWER #13

A is PD all $\lambda_i > 0$

Positive eigenvalues = PD

Why?

$x^T Ax = \Sigma\lambda_i y_i^2$ in eigenbasis

Use: When eigenvalues easy to find

⚠ > not ≥ for PD!

Back

CARD #14

PD Test Method 2 (Sylvester)

Hint: Leading minors

PD TESTS

INTERMEDIATE

Front

ANSWER #14

A is PD all leading principal minors > 0

$M_1>0, M_2>0, \ldots, M_n>0$

Why?

Sylvester's criterion

Use: Small matrices (2×2, 3×3)

⚠ Need ALL minors > 0!

Back

CARD #15

First-order necessary condition?

Hint: Gradient

OPTIMALITY

BASIC

Front

ANSWER #15

Local extremum $\nabla f(x^*) = 0$

Flat slope at extrema

Why?

Fermat's rule

Use: Finding critical points

⚠ Necessary, not sufficient!

Back

CARD #16

Second-order sufficient condition for minimum?

Hint: Grad + Hess

OPTIMALITY

INTERMEDIATE □□

Front

ANSWER #16

$\nabla f(x^*)=0$ AND $\nabla^2 f(x^*)>0$ □ strict local min

Zero grad + PD Hessian

Why?

Parabola opening upward

Use: Proving local minimum

⚠ Need BOTH conditions!

Back

CARD #17

Hessian test: $\det(H)>0$, $H_{11}>0$ □ ?

Hint: Bowl shape

HESSIAN 2×2

INTERMEDIATE □□

Front

ANSWER #17

LOCAL MINIMUM

Positive det + positive diagonal

Why?

Both eigenvalues > 0

Use: 2×2 critical point classification

⚠ Need BOTH conditions!

Back

CARD #18

Hessian test: $\det(H) < 0$ □ ?

Hint: Mountain pass

HESSIAN 2×2

INTERMEDIATE □□

Front

ANSWER #18

SADDLE POINT

Negative det = opposite curvatures

Why?

Eigenvalues have different signs

Use: Identifying saddles

⚠ Just need det<0!

Back

CARD #19

Characteristic equation?

Hint: Find λ

EIGENVALUES

BASIC □

Front

ANSWER #19

$\det(A - \lambda I) = 0$

Determinant equals zero

Why?

$Av=\lambda v$ □ $(A-\lambda I)v=0$
Non-trivial v □ singular

Use: Finding eigenvalues

⚠ Set DET = 0, not Av = 0!

Back

CARD #20

Convexity definition?

Hint: Line segment

CONVEXITY

INTERMEDIATE □□

Front

ANSWER #20

$f(\theta x+(1-\theta)y) \leq \theta f(x)+(1-\theta)f(y)$ for $\theta\in[0,1]$, all x,y

Chord above graph

Why?

Line segment stays above function

Use: Proving convexity

⚠ Must hold for ALL x,y!

Back

FLASHCARD SET - Page 5 of 5 | Cards 21-25

CARD #21

First-order convexity test?

Hint: Tangent below

CONVEXITY TEST

INTERMEDIATE □□

Front

ANSWER #21

f convex □ $f(y) \geq f(x) + \nabla f(x)^T(y-x)$

Function above all tangents

Why?

Linear approximation underestimates

Use: Checking convexity (differentiable)

⚠ For ALL x,y!

Back

CARD #22

Second-order convexity test?

Hint: Hessian

CONVEXITY TEST

INTERMEDIATE □□

Front

ANSWER #22

f convex □ $\nabla^2 f(x)$ □ 0 for all x

PSD Hessian everywhere

Why?

Positive curvature everywhere

Use: Twice differentiable f

⚠ Check ALL x in domain!

Back

CARD #23

Convex + $\nabla f(x^*)=0$ □ ?

Hint: Global property

KEY THEOREM

ADVANCED □□□

Front

ANSWER #23

□ x^* is GLOBAL MINIMUM

Convex: stationary = global min

Why?

No local minima in convex functions

Use: Proving global optimality

⚠ Only works if f is CONVEX!

Back

CARD #24

Cauchy-Schwarz inequality?

Hint: DPL

INEQUALITY

BASIC □

Front

ANSWER #24

$|x^T y| \leq ||x||_2 \cdot ||y||_2$

DPL: Dot ≤ Product of Lengths

Why?

From $|\cos(\theta)| \leq 1$

Use: Bounding inner products

⚠ Don't forget absolute value!

Back

CARD #25

Angle between vectors x, y?

Hint: Cosine formula

INNER PRODUCT

BASIC □

Front

ANSWER #25

$\cos(\theta) = x^T y / (||x||\cdot||y||)$

Dot over product of lengths

Why?

From geometric definition

Use: Finding angles

⚠ Normalize with BOTH norms!

Back

SAMPLE FLASHCARD: Gradient of x^T x

QUESTION SIDE

What is

$$\nabla(x^T x) = ?$$

Hint: Think "SDD"

Category: GRADIENT FORMULAS

Difficulty: BASIC

Card 1 of 25 | Most Important Formula!

ANSWER SIDE

$$\nabla(x^T x) = 2x$$

Mnemonic: SDD

Self Dot Doubles

Why?

$x^T x = x_1^2 + x_2^2 + ... + x_n^2$

$d/dx_1 = 2x_1,\ d/dx_2 = 2x_2,\ ...$

$\therefore \nabla = [2x_1, 2x_2, ..., 2x_n]^T = 2x$

Use when:

Finding ∇f when $f(x) = ||x||_2^2$

⚠ Don't forget the 2!

QUICK REFERENCE SHEET - All 25 Formulas at a Glance

GRADIENT FORMULAS (Cards 1-5)

#1	*SDD*	$\nabla(x^T x) = 2x$
#2	*LKC*	$\nabla(a^T x) = a$
#3	*QAT*	$\nabla(x^T Ax) = (A+A^T)x$
#4	*QAT*	If $A=A^T$: $\nabla(x^T Ax) = 2Ax$
#5	*NSDA*	$\nabla(\|\|Ax\|\|^2) = 2A^T Ax$

VECTOR NORMS (Cards 6-9)

#6	*LAA*	$\|\|x\|\|_1 = \Sigma\|x_i\|$
#7	*SSR*	$\|\|x\|\|_2 = \sqrt{(\Sigma x_i^2)}$
#8	*MM*	$\|\|x\|\|_\infty = \max\|x_i\|$
#9	--	$\|\|x\|\|_\infty \leq \|\|x\|\|_2 \leq \|\|x\|\|_1$

MATRIX PROPERTIES (Cards 10-12)

#10	*RROS*	$(AB)^T = B^T A^T$
#11	--	$A^T A$ always PSD
#12	*SPD*	$A=A^T$ AND $x^T Ax>0$

PD TESTS (Cards 13-14)

#13	*λ*	PD □ all $\lambda_i > 0$
#14	*Syl*	PD □ all minors > 0

OPTIMALITY (Cards 15-16)

#15	*1°*	$\nabla f(x^*) = 0$ (necessary)
#16	*2°*	$\nabla f=0$ & $H>0$ □ min

MNEMONIC KEY:

SDD=Self Dot Doubles | LKC=Linear Keeps Constant | QAT=Quadratic Adds Transpose

NSDA=Norm Squared Doubles A^T A | LAA=L_1 Adds Absolutes | SSR=Square Sum Root

MM=Max/Maximum | RROS=Reverse Roll Order Switches | SPD=Symmetric Positive Definite

DPL=Dot ≤ Product Lengths | Syl=Sylvester | Cvx=Convex | Thm=Theorem

#17	*H*	$\det(H)>0$, $H_{11}>0$ □ MIN
#18	*H*	$\det(H)<0$ □ SADDLE
#19	*λ*	$\det(A-\lambda I) = 0$

#20	*Cvx*	$f(\theta x+(1-\theta)y)\leq\theta f(x)+(1-\theta)f(y)$
#21	*1°*	$f(y)\geq f(x)+\nabla f^T(y-x)$
#22	*2°*	$\nabla^2 f(x)$ □ 0 $\forall x$

#23	*Thm*	Cvx+$\nabla f=0$ □ global min
#24	*DPL*	$\|x^T y\| \leq \|\|x\|\|\cdot\|\|y\|\|$
#25	*θ*	$\cos(\theta)=x^T y/(\|\|x\|\|\cdot\|\|y\|\|)$