OPTIMIZATION FORMULA SHEET - Quick Reference

GRADIENT FORMULAS

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

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

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

If $A = A^T$: $\nabla(x^T A x) = 2Ax$

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

VECTOR NORMS

$\|x\|_1 = \sum_i |x_i|$

$\|x\|_2 = \sqrt{\sum_i x_i^2}$

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

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

MATRIX PROPERTIES

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

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

$A^T A$ is always PSD

$A^T A$ is symmetric

$x^T A y = y^T A^T x$ if scalar

POSITIVE DEFINITE

A is PD $\Leftrightarrow x^T A x > 0 \ \forall x \neq 0$

A is PD $\Leftrightarrow$ all $\lambda_i > 0$

A is PD $\Leftrightarrow$ all minors > 0

SPD = Symmetric + PD

OPTIMALITY CONDITIONS

1st order necessary: $\nabla f(x^*) = 0$

2nd order necessary: $\nabla^2 f(x^*) \succcurlyeq 0$

2nd order sufficient:

$\nabla f(x^*) = 0$ & $\nabla^2 f(x^*) \succ 0$

$\Rightarrow$ strict local min

HESSIAN TEST (2×2)

$\det(H) > 0, H_{11} > 0 \Rightarrow$ MIN

$\det(H) > 0, H_{11} < 0 \Rightarrow$ MAX

$\det(H) < 0 \Rightarrow$ SADDLE

$\det(H) = 0 \Rightarrow$ INCONCLUSIVE

CONVEXITY TESTS

Definition: $f(\theta x + (1-\theta)y)$

$\leq \theta f(x) + (1-\theta) f(y)$

1st order: $f(y) \geq f(x) + \nabla f(x)^T (y - x)$

2nd order: $\nabla^2 f(x) \succcurlyeq 0 \ \forall x$

KEY THEOREMS

Convex + $\nabla f = 0 \Rightarrow$ global min

Strictly convex $\Rightarrow$ ≤ 1 global min

Convex: local min = global min

Cauchy-Schwarz: $|x^T y| \leq \|x\| \|y\|$

EIGENVALUES

$Av = \lambda v$ (definition)

$\det(A - \lambda I) = 0$ (char. eq.)

Symmetric A: real λ, orthogonal v

A PD $\Leftrightarrow$ all $\lambda_i > 0$

TAYLOR EXPANSION

$f(x+h) \approx f(x)$

$+\nabla f(x)^T h$

$+\frac{1}{2} h^T \nabla^2 f(x) h$

0th + 1st + 2nd order

INNER PRODUCT

$x^T y = \sum_i x_i y_i$ (dot product)

$x^T x = \|x\|_2^2$

$\cos(\theta) = \frac{x^T y}{\|x\| \|y\|}$

$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2x^T y$

QUICK MNEMONICS

SDD: Self Dot Doubles

QAT: Quadratic Adds Transpose

LAA: L_1 Adds Absolutes

SSR: Square Sum Root (L_2)

SPD: Symmetric Positive Definite

DPL: Dot ≤ Product of Lengths