
Matrix Calculus Cheatsheet

Synthetic formula-heavy sample for layout regions, math symbols, and compact prose

Core Identities

1. $d/dx (a \odot x) = a$
2. $d/dx (x \odot Ax) = (A + A^T)x$
3. $\nabla_x \|Ax - b\|^2 = 2A^T(Ax - b)$
4. $\partial/\partial X \operatorname{tr}(AXB) = A^T B^T$
5. $\operatorname{softmax}(z_i) = \exp(z_i) / \sum_j \exp(z_j)$
6. $\operatorname{cross_entropy}(y, p) = -\sum_i y_i \log p_i$

Worked Example

Given $f(x) = \|Wx - y\|^2 + \lambda \|x\|^2$

Gradient: $\nabla f(x) = 2W^T(Wx - y) + 2\lambda x$

Normal equation: $(W^T W + \lambda I)x = W^T y$

If $W = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $y = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, solve by inversion or QR factorization.