1. Algebra:

a. $(a \pm b)^2 = a^2 \pm 2ab + b^2$

b. $(a+b)(a-b) = a^2 - b^2$

c. $(a \pm b)^3 = a^3 \pm 3a^2 b + 3ab^2 \pm b^3$

d. $a^3 - b^3 = (a-b)(a^2 + ab + b^2)$

e. $ax^2 + bx + c = a \cdot (x - x_1)(x - x_2)$

where $x_{1,2} = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ and $a \neq 0$.

2. Powers and Logarithms:

a. $a^x \cdot b^x = (a \cdot b)^x$

b. $\dfrac{a^x}{b^x} = \left(\dfrac{a}{b}\right)^x$

c. $a^x \cdot a^y = a^{x+y}$

d. $\dfrac{a^x}{a^y} = a^{x-y}$

e. $\log_a y = x \Leftrightarrow a^x = y$ for $a > 0, a \neq 1$

f. $\log_a a^x = x$

g. $a^{\log_a y} = y$

h. $\log_a (x \cdot y) = \log_a x + \log_a y$

i. $\log_a \dfrac{x}{y} = \log_a x - \log_a y$

j. $\log_a x^y = y \cdot \log_a x$

k. $\log_a x = \frac{\log_b x}{\log_b a}$

l $(a^x)^y = a^{x \cdot y}$ $\quad a^{-x} = \frac{1}{a^x}$ $\quad a^{\frac{x}{y}} = \sqrt[y]{a^x}$ $\quad (y \neq 0 \;\; a \neq 0)$

3. Derivatives

a. $(f \pm g)'(x) = f'(x) \pm g'(x)$

b. $(c \cdot f)'(x) = c \cdot f'(x)$

c. $(f \cdot g)'(x) = f'(x) g(x) + f(x) g'(x)$

d. $\left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x) g(x) - f(x) g'(x)}{g^2(x)}$

e. $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$

f. $[f^{-1}(y)]' = \dfrac{1}{f'(f^{-1}(y))}$

g. $(\sqrt{x})' = \frac{1}{2\sqrt{x}}$

h. $(\log_a x)' = \dfrac{1}{x \cdot \ln a}$ $\quad (\cot x)' = -\dfrac{1}{\sin^2 x}$

4. Gradient

$$\nabla f(x_1, x_2) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}\right)$$

partial derivative:

$$D_u f(x_1, x_2) = \lim_{s \to 0} \frac{f(x_1 + s u_1, x_2 + s u_2) - f(x_1, x_2)}{s}$$

$$= \left(\frac{df}{ds}\right)_u = \nabla f(x_1, x_2) \cdot u$$

Directional derivative (נגזרת כיוונית):

$$D_u f(x_1, x_2) = \nabla f(x_1, x_2) \cdot \vec{u} = |\nabla f||\vec{u}| \cos\beta = |\nabla f| \cos\beta$$

5. Series

The sum of an **arithmetic series**:

$$S_n = \frac{(a_1 + a_n) n}{2} = \frac{[2a_1 + d(n-1)] n}{2}$$

The sum of a **geometric series**:

$$S_n = \frac{a_1 (q^n - 1)}{q - 1}$$

6. Cost Function

Regression: $J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} \left(\theta \cdot x^{(i)} - y^{(i)}\right)^2$

Perceptron : $E[\vec{w}] = \frac{1}{2} \sum_{d \in D} \left(sgn(\vec{w} \cdot \vec{x}^{(d)}) - t_d\right)$

LMS: $E[\vec{w}] = \frac{1}{2} \left[\sum_{d \in D} \left(\vec{w} \cdot \vec{x}^{(d)} - 1\right)^2 + \left(\vec{w} \cdot \vec{x}^{(d)} + 1\right)^2\right]$

SVM Optimization problem (Maximization):

$$max_{w_0, \underline{w}}\, b \text{ s.t. } \frac{t_d \left(\sum_{i=1}^{n} w_i x_i^{(d)} + w_0\right)}{\| w \|} \geq b.$$

7. Goodness of split

$$\Delta\varphi(S, A) = \varphi(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} \varphi(S_v)$$

8. Probability

$$E(X) = \sum_i x_i * P(X = x_i)$$

$$E(X) = \int_{-\infty}^{\infty} x * f(x) dx$$

Variance:

$$var(X) = E[(X - E(X)^2] = cov(X, X)$$

Covariance:

$$cov(X, Y) = E[(X - E(X))(Y - E(Y))]$$

Bayes:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) * P(B|A)}{P(B)}$$

Minkowski Matric

$$L_k(a, b) = \left(\sum_{i=1}^{d} |a_i - b_i|^k\right)^{\frac{1}{k}} \quad k \geq 1$$

Manhattan Distance

$$L_1 = |a - b|$$

Euclidean Distance

$$L_2 = \left(\sum_{i=1}^{d} |a_i - b_i|^2\right)^{\frac{1}{2}}$$

Infinity Norm

$$L_\infty = max(|a_i - b_i|)$$

Pseudo- Inverse: $pinv(X) = (X^T X)^{-1} X^T$

Inverse matrix to two- dimension matrix:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$