

140 Machine Learning Formulas

Rubens Zimbres - 2017 - version 2

English version

.PDF 18 pages

rubenzmm@gmail.com

<http://github.com/RubensZimbres>

NAÏVE BAYES

$$P(a|c) = \frac{P(c|a).P(a)}{P(c)}$$

MIXTURE MODELS

$$P(B) = P(B|A).P(A)$$

BAYES OPTIMAL CLASSIFIER

$$\arg \max \sum P(x|T).P(T|D)$$

**MIXTURE OF GAUSSIANS
ANOMALY DETECTION**

$$P(x|\bar{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2\right]$$

NAÏVE BAYES CLASSIFIER

$$\arg \max P(Spo|Tot) \cdot \prod P(Soc|Spo)$$

$$Z_{CS} = \frac{N_A C_B + N_B C_A}{N_A + N_B}$$

$$P(Z_{cs}) \rightarrow 0.50$$

BAYES MAP (maximum a posteriori)

$$h_{MAP} = \arg \max P(c|a).P(a)$$

EM ALGORITHM

$$E \text{ step } P(\bar{x}|x) = \frac{P(\bar{x}).P(x|\bar{x})}{\sum P(x).P(\bar{x})}$$

MAXIMUM LIKELIHOOD

$$h_{ML} = \arg \max P(c|a)$$

$$M \text{ step } P(x') = \frac{\sum P(\bar{x}|x)}{n}$$

TOTAL PROBABILITY

$$TotalP(B) = P(B|A).P(A)$$

E step $P(\bar{x}|x) = \text{Assign value}$

$$M \text{ step } P(x') = P(B = 1|A = 1, C = 0)$$

LAPLACE ESTIMATE (small samples)

$$\frac{d}{dx} \frac{f(x)}{g(x)} = \frac{f'(x)g(x) + f(x).g'(x)}{g(x)^2}$$

$$P(A) = \frac{A + 0.5}{A + B + 1}$$

$$\frac{d}{dx} 2f(x) = 2 \frac{d}{dx} f(x)$$

BAYESIAN NETWORKS

tuples \neg for $y = 0 \wedge y = 1$

$$\frac{d}{dx} f(x) + g(x) = \frac{d}{dx} f(x) + \frac{d}{dx} g(x)$$

LIMITS

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

$$\frac{d}{dx} f(x) + 2g(x) = \frac{d}{dx} f(x) + 2 \frac{d}{dx} g(x)$$

$$h = \Delta x = x' - x$$

CHAIN RULE

$$\frac{d}{dx} g(f(x)) = g'(f(x)).f'(x)$$

DERIVATIVES

solve $f(x)$ apply in $g'(x)$

$$\frac{\partial}{\partial x} x^n = n \cdot x^{n-1}$$

VARIANCE

$$\frac{\partial}{\partial x} y^n = \frac{\partial y^n}{\partial y} \cdot \frac{\partial y}{\partial x}$$

$$Var = \frac{\sum(x - \bar{x})^2}{n - 1}$$

PRODUCT RULE

$$\frac{d}{dx} f(x) \cdot g(x) = f'(x)g(x) + f(x).g'(x)$$

STANDARD DEVIATION

$$\sqrt{Variance}$$

COVARIANCE

$$Cov = \frac{(x - \bar{x}).(y - \bar{y})}{n - 1}$$

SUM OF SQUARED ERRORS

$$E|\vec{w}| = \frac{\sum(\hat{y} - y)^2}{2}$$

COST FUNCTION**CONFIDENCE INTERVAL**

$$x \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

$$J(\theta_j) := \theta_j - \eta \cdot \frac{\sum(\hat{y} - y)^2}{2}$$

CHI SQUARED

$$Chi = \frac{(\hat{y} - y)^2}{\sqrt{y}} = \frac{\delta^2}{\sqrt{y}}$$

NUMBER OF EXAMPLES

$$m \geq \frac{\log(N_H) + \log(\frac{1}{\delta})}{\epsilon}$$

$$\text{where } \epsilon = \frac{\hat{y}}{y} \wedge \delta = y - \hat{y}$$

R SQUARED

$$R^2 = \frac{n \sum xy - \sum x \cdot \sum y}{\sqrt{(n \sum x^2 - (\sum x)^2) \cdot (n \sum y^2 - (\sum y)^2)}}$$

MARKOV CHAINS

$$P^{t+1}(X = x) = \sum_x P^t \cdot (X = x) \cdot T(x \rightarrow x)$$

LOSS

$$Loss = Bias^2 + Variance^2 + Noise$$

K NEAREST NEIGHBOR

$$\hat{f}(x) \leftarrow \frac{\sum f(x)}{k}$$

$$DE(x_i, x_j) = \sqrt{(x_i - x_j)^2 + (y_{xi} - y_{xj})^2}$$

WEIGHTED NEAREST NEIGHBOR

$$f(x) = \sum \frac{f(x)}{D(x_1x_2)^2} \cdot \sum D(x_1x_2)^2$$

t-SNE

$$Condit. Prob = \frac{\exp\left(-\frac{||x_i - x_j||^2}{2\sigma^2}\right)}{\sum \exp\left(-\frac{||x_i - x_k||^2}{2\sigma^2}\right)}$$

$$Condit. Prob = \frac{\exp\left(-\frac{||y_i - y_j||^2}{2\sigma^2}\right)}{\sum \exp\left(-\frac{||y_i - y_k||^2}{2\sigma^2}\right)}$$

$$Perplexity^{(P_i)} = 2^{H(P_i)}$$

PRINCIPAL COMPONENTS ANALYSIS

$$x' = x - \bar{x}$$

$$Eigenvalue = [A] - \lambda I$$

$$Eigenvector = Engenvalue. [A]$$

$$f(x) = Eigenvector^T. [x_{i1} \dots x_{jn}]$$

where:

$$H(P_i) = - \sum_j p_{j|i} \log_2 P_{j|i}$$

LINEAR REGRESSION

$$m_1 = \frac{\sum x_2^2 \sum x_1 y - \sum x_1 x_2 \sum x_2 y}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2}$$

$$b = \bar{y} - m_1 \bar{x}_1 - m_2 \bar{x}_2$$

$$\bar{x}_1 \sim \bar{x}_2 \neq \bar{x}_1' \sim \bar{x}_2'$$

$$f(x) = \sum_{i=1}^n m_i x_i + b$$

$$mx + b = \frac{p}{1-p}$$

$$A = (X^T.X)^{-1}.X^T.Y$$

$$\text{where } A = \begin{bmatrix} b \\ m \end{bmatrix}$$

$$P(a|c) = \frac{mx+b}{mx+b+1}$$

DECISION TREES

LOGISTIC REGRESSION

$$Entropy = \sum_{v=0}^1 -P_v \log(P_v)$$

$$Odds\ Ratio = \log\left(\frac{P}{1-P}\right) = mx + b$$

$$InfoGain = P_+ \cdot [-P_{+t} \cdot \log(P_{+t}) - P_{+(t-1)} \cdot \log(P_{+(t-1)})]$$

$$\left(\frac{P}{1-P}\right) = e^{mx+b}$$

RULE INDUCTION

$$Gain = P \cdot [(-P_{t-1} \cdot \log(P)) - (-P_t \cdot \log(P))]$$

$$J(\theta) = -\frac{\sum y \cdot \log(\hat{y}) + (1-y) \cdot \log(1-\hat{y})}{n}$$

RULE VOTE

$$\text{where } \hat{y} = \frac{1}{1 + e^{mx+b}}$$

$$Weight = accuracy \cdot coverage$$

for $y = 0 \wedge y = 1$

ENTROPY

$$-2LL \rightarrow 0$$

$$H(A) = -\sum P(A) \cdot \log P(A)$$

JOINT ENTROPY

$$H(A, B) = - \sum P(A, B) \cdot \log P(A, B)$$

SIMILARITY

$$w_{ij} = \frac{\sum_k (R_{ix} - \bar{R}_i) \cdot (R_{jk} - \bar{R}_j)}{\sqrt{\sum_k (R_{ix} - \bar{R}_i)^2 \cdot (R_{jk} - \bar{R}_j)^2}}$$

CONDITIONAL ENTROPY

$$H(A|B) = - \sum P(A, B) \cdot \log P(A|B)$$

CONTENT-BASED RECOMMENDATION

$$\text{Rating} = \sum_{i=1}^{\text{class}} \sum_{j=1}^m x_i y_j$$

MUTUAL INFORMATION

$$I(A, B) = H(A) - H(A|B)$$

COLLABORATIVE FILTERING

$$\hat{R}_{ik} = \bar{R}_i + \alpha \cdot \sum \left((R_{jk} - \bar{R}_j) \cdot \frac{\sum_k (R_{ix} - \bar{R}_i) \cdot (R_{jk} - \bar{R}_j)}{\sqrt{\sum_k (R_{ix} - \bar{R}_i)^2 \cdot (R_{jk} - \bar{R}_j)^2}} \right)$$

EIGENVECTOR CENTRALITY = PAGE RANK

$$PR(A) = \frac{1-d}{n} - d \left(\frac{PR(B)}{Out(B)} + \frac{PR(n)}{Out(n)} \right)$$

where $d=1$ few connections

RATING

$$\hat{R} = \bar{R}_i + \alpha \sum w_j \cdot (R_{jk} - \bar{R}_j)$$

BATCH GRADIENT DESCENT

$$J(\theta_j) := \theta_j \pm \eta \cdot \frac{\sum (\hat{y} - y)^2 \cdot x}{2n}$$

STOCHASTIC GRADIENT DESCENT

$$J(\theta_j) := \theta_j \pm \eta \cdot (\hat{y} - y)^2 \cdot x$$

LOSS

$$\textit{Loss} = \frac{\sum H(S(f(x), f(x)))}{N}$$

NEURAL NETWORKS

$$f(x) = o = w_0 + \sum_{i=1}^n w_i x_i$$

L2 REGULARIZATION

$$w \leftarrow w - \left(\eta \cdot \delta \cdot x + \frac{\lambda \cdot w^2}{2} \right)$$

LOGIT

$$\log(\textit{odds}) = wx + b = \log\left(\frac{p}{1-p}\right)$$

SIGMOID

$$\frac{1}{1 + e^{-(wx+b)}}$$

SOFTMAX NORMALIZATION

$$S(f(x)) = \frac{e^{wx+b}}{\sum e^{wx+b}}$$

RADIAL BASIS FUNCTION

$$h(x) = e^{\left(-\frac{(x-c)^2}{r^2}\right)}$$

CROSS ENTROPY

$$H(S(f(x), f(x))) = - \sum f(x) \cdot \log S(f(x))$$

PERCEPTRON

$$f(x) = \textit{sign} \left[\sum_{i=1}^n w_j x_{ij} \right]$$

PERCEPTRON TRAINING

$$w_i \leftarrow w_i + \Delta w_i$$

$$\Delta w_i = \eta \cdot (t - o) \cdot x$$

$$w_{ij} \leftarrow w_{ij} + \eta_{xi} \cdot \delta_h \cdot x_{ij}$$

$$w_1 = 1 + (t - o_j)$$

$$\Delta w_{ij}(n) = \eta \cdot \delta_k \cdot x_{ij} + M \cdot \Delta w_{ij}(n-1)$$

ERROR FOR A SIGMOID

$$\epsilon = \sum (t - o) \cdot o \cdot (1 - o) \cdot x$$

where $M=momentum$

NEURAL NETWORKS COST FUNCTION

$$J_\theta = \frac{\sum_{i=1}^n \sum_{k=1}^K t_k \cdot \log(o) + (1-t) \cdot \log(1-o)}{N} + \frac{\lambda \sum_{l=1}^{L-1} \sum_{i=1}^{N_l} \sum_{j=1}^{N_{l+1}} \theta_{ji}^2}{2N}$$

AVOID OVERFIT NEURAL NETWORKS L2

$$\sum \vec{w} = \frac{\sum_{d \in D} \sum_{k \in Y} (t - o)^2}{2} + F \cdot \sum w_{ij}^2$$

MOMENTUM Y

$$\theta = \theta - (\gamma v_{t-1} + \eta \cdot \nabla J(\theta))$$

where $F=penalty$

BACKPROPAGATION

$$\delta_k = o_k \cdot (1 - o_k) \cdot (t - o_k)$$

NESTEROV

$$\theta = \theta - (\gamma v_{t-1} + \eta \cdot \nabla J(\theta - \gamma v_{t-1}))$$

$$\delta_h = o_h \cdot (1 - o_h) \cdot \sum w_{jk} \delta_k$$

ADAGRAD

$$\theta = \theta - \frac{\eta}{\sqrt{SSG_{diag} + \epsilon}} \cdot \nabla J(\theta)$$

RESTRICTED BOLTZMANN MACHINES

ADADELTA

$$\theta = \theta - \frac{RMS[\Delta\theta]_{t-1}}{RMS\nabla J(\theta)}$$

$$RMS[\Delta\theta] = \sqrt{E[\Delta\theta^2] + \epsilon}$$

$$E(v, h) = - \sum v_i h_j w_{ij}$$

where v = binary state visible

h = binary state hidden

RMSprop

$$\theta = \theta - \frac{\eta}{\sqrt{E[g^2] + \epsilon}} \cdot \nabla J(\theta)$$

$$p(v, h) = \frac{e^{-E(v, h)}}{\sum_{u,g} e^{-E(u, g)}}$$

$$p(v) = \frac{\sum_h e^{-E(v, h)}}{\sum_{u,g} e^{-E(u, g)}}$$

ADAM

$$\theta = \theta - \frac{\eta}{\sqrt{\hat{v}} + \epsilon} \cdot \hat{m}$$

$$\frac{\partial}{\partial w_{ij}} \log p(v) = < v_i \ h_j >^0 - < v_i \ h_j >^\infty$$

$$\hat{m} = \frac{\beta_1 m_{t-1} + (1 - \beta_1) \cdot \nabla J(\theta)}{1 - \beta_1}$$

$$\Delta w_{ij} = \eta \cdot \frac{\partial}{\partial w_{ij}} \log p(v)$$

$$\Delta w_{ij} = \eta \cdot (< v_i \ h_j >^0 - < v_i \ h_j >^1)$$

$$\hat{v} = \frac{\beta_2 v_{t-1} + (1 - \beta_2) \cdot \nabla J(\theta)^2}{1 - \beta_2}$$

SUPPORT VECTOR MACHINES

$$f(x) = sign[\lambda \cdot y \cdot K(x_i \cdot x_j)]$$

$$K(x_i \cdot x_j) = \exp\left[-\sqrt{\frac{(x_i - x_j)^2 + (y_i - y_j)^2}{width_{hist}}}\right]$$

$$\lambda \rightarrow \nabla L = 0$$

$$y = 1 \wedge y = -1$$

$$DotProduct = \overrightarrow{x_1} \cdot cos\theta$$

$$\cos^2\theta + \sin^2\theta = 1$$

$$\sin\theta = \frac{\sqrt{(x_i - x_j)^2 + (y_{xi} - y_{xj})^2}}{\overrightarrow{x_2}}$$

$$(x_1 \cdot x_2) = \sqrt{(x_1^2 + y_1^2) \cdot \left(1 - \frac{(x_1 - x_2)^2 + (y_1 - y_2)^2}{x_2^2 + y_2^2}\right)}$$

SUPPORT VECTOR REGRESSION

$$\hat{Y} = w \cdot \langle x_i \cdot x_j \rangle + b$$

$$y - (w \cdot \langle x_i \cdot x_j \rangle + b) \leq \varepsilon$$

$$w \cdot \langle x_i \cdot x_j \rangle + b - y \leq \varepsilon$$

RIDGE REGRESSION - REGULARIZATION

$$m := m - \frac{\sum(\hat{y} - y)^2}{N} - \frac{\lambda \cdot m}{N}$$

$$y = \lambda \cdot mx + b - \frac{\lambda}{N}$$

LASSO REGRESSION - REGULARIZATION

$$b := \frac{\sum(\hat{y} - y)^2}{N} + \frac{\lambda \cdot b}{N}$$

$$m \rightarrow 0$$

$$y = mx + \lambda \cdot b + \frac{\lambda}{N}$$

SKEWNESS

Skewness < 1

KOLMOGOROV SMIRNOV

Normal sig > .005

NON PARAMETRIC

T test = Normal
Test U Mann Whitney sig < .05

TOLERANCE

Tolerance > .1

$$Tolerance = \frac{1}{VIF}$$

CRONBACH

> .60 .70

VARIANCE INFLATION FACTOR

VIF <10

MEDIAN

$$\frac{Max - Min}{2}$$

ENTER METHOD

+ 15 cases / Variable

t TEST

$$t = \frac{(x_1 - x_2) - (\mu_1 - \mu_2)}{(x_1 - x_2)}$$

STEPWISE METHOD

Difference significant sig < .05

+ 50 cases / Variable

t TEST 2 SAMPLES

Levene Variância

VARIABLE SELECTION

F Test = 47 sig < .05

ANOVA + 3

$$F = \frac{Variance\ between\ groups}{Variância\ inside\ group}$$

Sig < .05

MISSING DATA

Delete if > 15%

DISCRIMINANT ANALYSIS

Box M sig < .05 reject H0

Wilk's Lambda sig < .05

$$\bar{x}_1 \sim \bar{x}_2 \neq \bar{x}_1' \sim \bar{x}_2'$$

$$P(x|\bar{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x - \bar{x}}{\sigma}\right)^2\right]$$

$$Z_{CS} = \frac{N_A C_B + N_B C_A}{N_A + N_B}$$

TRANSFORMATION OK

$$\frac{\bar{x}}{\sigma} < 4$$

MULTICOLLINEARITY

Correlation > .90

VIF < 10

Tolerance > .1

SUM OF SQUARES (explain)

$$F_{ratio} = \frac{SS_{regression} \cdot (N - coef)}{(coef - 1) \cdot SS_{residuals}}$$

ERROR MARGIN

$$1.96 \frac{\sigma}{\sqrt{N}}$$

STANDARD ERROR ESTIMATE (SEE)

$$SEE = \sqrt{\frac{\text{SumSquaredErrors}}{n - 2}}$$

ACCURACY

Confidence Interval ~ P value

HYPOTHESES TESTING

P value < .05

$$SEE = \sqrt{\frac{\sum(\hat{y} - y)^2}{n - 2}}$$

MAHALANOBIS DISTANCE

same variable

$$M = \sqrt{\frac{(x_1 - \bar{x}_1)^2}{\sigma^2}}$$

$$Q_{s,a} = R_s + \delta \max_{s'} \sum_s T(s, a, s'). \max_{a'} Q(s', a')$$

$$\hat{Q}_{s,a} \leftarrow \eta R_s + \delta \max_a Q(s', a')$$

ARIMA ~ NPV

MANHATTAN DISTANCE L

$$Manh = |x_1 - x_2| + |y_1 - y_2|$$

$$B^n Y_t = Y_{t-n}$$
 (Backward Shift Operator)

$$B^2 Y = B(BY_t) = B(Y_{t-1}) = Y_{t-2}$$

NET PRESENT VALUE

$$P_t = P_0 \cdot \theta^t$$

ARIMA(1,1,1):

$$P_0 = P_t \cdot \theta^{-t}$$

AR = number autoregressive terms

$$NPV = \text{investment} + \sum_{t=1}^N \frac{capital}{(1 + rate)^t}$$

B=number non-seasonal needed for stationary

$$NPV=0$$
 (IRR)

MA=number lagged errors

$$(1 - \phi_1 B)(1 - B)Y_t = (1 - \theta_1 B)e_t$$

where $(1 - \phi_1 B)$ =AR (Autoregression)

and $(1 - \theta_1 B)$ =MA (Mean Average)

MARKOV DECISION PROCESS

$$U_s = R_s + \delta \max_a \sum_s T(s, a, s'). U(s')$$

and $e=noise$

$$\pi_s = \operatorname{argmax}_a \sum_s T(s, a, s'). U(s')$$

PROBABILITY (coins)

$$P(a) = \frac{P(a)}{P(A)}$$

FREQUENTIST

$$\lim_{n \rightarrow \infty} = \frac{m}{n} = \frac{\text{sucessos}}{\text{todas possibilidades}} = \frac{\text{eventos}}{\text{espaço amostral}}$$

MARGINAL PROBABILITY

$$P(a) = \frac{P(A = a)}{\sum P(A)}$$

AXIOMATIC

$$P(A) \geq 0$$

$$\sum P(A, B, C) = 1$$

PROBABILITY A and B

$$P(A \text{ e } B) = \frac{P(A \cap B)}{P(B)}$$

PROBABILITY THEOREMS**JOIN = A or B**

$$P(A \cup B)_{\text{EXCLUDENT}} = P(A) + P(B)$$

$$P(A \cup B)_{\text{NOT EXCLUDENT}} = P(A) + P(B) - P(A \cap B)$$

CONDITIONAL PROBABILITY

$$P(A|B)_{\text{INDEPENDENTS}} = P(A)$$

$$\begin{aligned} P(A \cup B \cup C)_{\text{NOT EXCLUDENT}} \\ = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) - P(A \cap B \cap C) \end{aligned}$$

COMPLEMENTARY EVENT

$$P(\tilde{A}) = 1 - P(A)$$

BAYES (52 cards , cancer)

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

BINOMIAL DISTRIBUTION (0,1 success)

$$P(D) = \binom{\text{sample space}}{\text{sucess}} \cdot P(s)^s \cdot (1 - P(s))^{N-s}$$

$$P(D) = \binom{\text{sample space}}{\text{sucess}} \cdot P(s)^s \cdot (P(\bar{s}))^{N-s}$$

CHAIN RULE

$$P(D) = \frac{c!}{a!(c-a)!} \cdot P(a)^a \cdot (1 - P(a))^{c-a}$$

$$\int f(x) + g(x). dx = \int f(x). dx + \int g(x). d(x)$$

TOTAL PROBABILITY (jars)

$$P(B) = \sum P(A \cap B) = \sum P(A) \cdot P(B|A)$$

INTEGRATION

$$\sum f'(x) \cdot \Delta x \underset{N \rightarrow \infty}{=} 0$$

PROBABILITY k SUCCESS in n TRIALS

$$P(k \text{ in } n) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

INTEGRALS

$$\int_a^b F(b) - F(a)$$

DIFFERENTIATION

$$\lim_{n \rightarrow \infty} \frac{f(a + \Delta x) - f(a)}{\Delta x}$$

$$\int_1^2 x^2 dx = \frac{1}{3} x^3 = \frac{1}{3} 2^3 - \frac{1}{3} 1^3$$

LINEAR ALGEBRA**ADDITION**

$$\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 2 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 9 & 6 \end{bmatrix}$$

PRODUCT RULE

$$\int c \cdot f'(x). dx = c \int f'(x). dx$$

SCALAR MULTIPLY

$$3 * \begin{bmatrix} 2 & 2 \\ 5 & 3 \end{bmatrix} = \begin{bmatrix} 6 & 6 \\ 15 & 9 \end{bmatrix}$$

MATRIX VECTOR MULTIPLICATION

Rows x Columns

IMPORTANT

x Vector: Column A = Rows B

A_{2,3} = 2a row x 3a column

$$A_{i,j} * B_{j,i} = C_{i,i}$$

$$\begin{bmatrix} 0 & 3 \\ 1 & 3 \\ 2 & 4 \end{bmatrix} * \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 9 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 5 \\ 0 & 3 & 2 \end{bmatrix} * \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 9 \\ 6 \end{bmatrix}$$

OR

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 5 \\ 0 & 3 & 2 \end{bmatrix} * \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = 1 * \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 2 * \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} + 0 * \begin{bmatrix} 3 \\ 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 9 \\ 6 \end{bmatrix}$$

PERMUTATION

LEFT=exchange rows

**x Matrix: Column A = Rows B
Rows A = Column B**

A_{2,1} = 2nd row x 1a column

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \end{bmatrix} * \begin{bmatrix} 0 & 3 \\ 1 & 3 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 8 & 24 \\ 14 & 37 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 0 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} * \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = [12 \quad 30 \quad 0]$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} * \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

RIGHT=exchange columns

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} * \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

IDENTITY

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

DIAGONAL

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

DETERMINANT

$$\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix} = 1.2 - 3.4 = -10$$

TRANSPOSE

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

DEMAND ELASTICITY

$$\rho = \frac{(Q_1 - Q_0)}{(Q_1 + Q_0)} \cdot \frac{(P_1 + P_0)}{(P_1 - P_0)}$$

PROPERTIES

Not commutative

$$A * B \neq B * A$$

Associative

$$A * B * C = A * (B * C)$$

Inverse (only squared)

$$A^{-1} \neq \frac{1}{A}$$

$$A^{-1} \cdot A = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = 1.5.9 + 4.8.3 + 7.2.6 - 7.5.3 - 1.8.6 - 4.2.9$$