150 Machine Learning Formulas

Rubens Zimbres – 2017

English version

.PDF 19 pages

rubenszmm@gmail.com

http://github.com/RubensZimbres

NAÏVE BAYES

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

BAYES OPTIMAL CLASSIFIER

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

NAÏVE BAYES CLASSIFIER

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

BAYES MAP (maximum a posteriori)

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

MAXIMUM LIKELIHOOD

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

TOTAL PROBABILITY

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

MIXTURE MODELS

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

MIXTURE OF GAUSSIANS ANOMALY DETECTION

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

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

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

EM ALGORITHM

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

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

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

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

LAPLACE ESTIMATE (small samples)

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

BAYESIAN NETWORKS

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

LIMITS

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

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

DERIVATIVES

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

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

PRODUCT RULE

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

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

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

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

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

CHAIN RULE

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

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

VARIANCE

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

STANDARD DEVIATION

$$\sqrt{Variance}$$

COVARIANCE

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

CONFIDENCE INTERVAL

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

CONFIDENCE INTERVAL ERROR

$$error \pm 1.96.\sqrt{\frac{error(1-error)}{N}}$$

CHI SQUARED

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

R SQUARED

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

LOSS

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

SUM OF SQUARED ERRORS

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

COST FUNCTION

$$J(\theta_j) \coloneqq \theta_j - \eta.\frac{\sum(\hat{y}-y)^2}{2}$$

GINI COEFFICIENT

$$Gini = \frac{N+1-2.\frac{\sum(N+1-x).y_i}{\sum y}}{N}$$

NUMBER OF EXAMPLES

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

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

MARKOV CHAINS

$$P^{t+1}(X = x) = \sum_{x} P^{t}.(X = x).T(x \to x)$$

K NEAREST NEIGHBOR

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

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

WEIGHTED NEAREST NEIGHBOR

$$f(x) = \sum \frac{f(x)}{D(x_1 x_2)^2} . \sum D(x_1 x_2)^2$$

PRINCIPAL COMPONENTS ANALYSIS

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

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

$$Eigenvector = Engenvalue.[A]$$

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

t-SNE

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

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

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

where:

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

COSINE DISTANCE

$$Cos = \frac{u.v}{\|u\|.\|v\|}$$

TF-IDF

$$w_{ij} = tf_{ij}.log\frac{N}{df_i}$$

LINEAR REGRESSION

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

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

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

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

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

LOGISTIC REGRESSION

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

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

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

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

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

$$-2LL \rightarrow 0$$

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

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

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

DECISION TREES

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

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

RULE INDUCTION

$$Gain = P.\left[(-P_{t-1}.\log(P)) - (-P_{t}.\log(P))\right]$$

RULE VOTE

Weight=accuracy . coverage

ENTROPY

$$H(A) = -\sum P(A).logP(A)$$

JOINT ENTROPY

$$H(A,B) = -\sum P(A,B).logP(A,B)$$

CONDITIONAL ENTROPY

$$H(A|B) = -\sum P(A,B).logP(A|B)$$

MUTUAL INFORMATION

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

EIGENVECTOR CENTRALITY = PAGE RANK

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

where d=1 few connections

RATING

$$\hat{R} = \bar{R}_{\iota} + \alpha \sum w_{j}.(R_{jk} - \bar{R}_{j})$$

SIMILARITY

$$w_{ij} = \frac{\sum_{k}(R_{ix} - \bar{R}_{i}).(R_{jk} - \bar{R}_{j})}{\sqrt{\sum_{k}(R_{ix} - \bar{R}_{i})^{2}.(R_{jk} - \bar{R}_{j})^{2}}}$$

CONTENT-BASED RECOMMENDATION

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

COLLABORATIVE FILTERING

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

BATCH GRADIENT DESCENT

$$J(\theta_j) \coloneqq \theta_j \pm \eta.\frac{\sum(\hat{y} - y)^2.x}{2n}$$

STOCHASTIC GRADIENT DESCENT

$$J(\theta_j) \coloneqq \theta_j \pm \eta.(\hat{y} - y)^2.x$$

NEURAL NETWORKS

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

LOGIT

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

SOFTMAX NORMALIZATION

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

CROSS ENTROPY

$$H(S(f(x), f(x)) = -\sum f(x).logS(f(x))$$

LOSS

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

L2 REGULARIZATION

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

SIGMOID

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

RADIAL BASIS FUNCTION

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

PERCEPTRON

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

PERCEPTRON TRAINING

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

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

ERROR FOR A SIGMOID

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

AVOID OVERFIT NEURAL NETWORKS L2

$$\sum \vec{w} = \frac{\sum_{d\epsilon D}\sum_{k\epsilon y}(t-o)^2}{2} + \mathrm{F}.\sum w_{ij}^2$$

where F=penalty

BACKPROPAGATION

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

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

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

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

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

where M=momentum

NEURAL NETWORKS COST FUNCTION

$$J_\theta = \frac{\sum_{i=1}^{n}\sum_{k=1}^{k} t_k.\log(o) + (1-t).\log(1-o)}{N} + \frac{\lambda \sum_{l=1}^{ii}\sum_{i=1}^{jj}\sum_{j=1}^{jj+1}\theta_{ji}^2}{2N}$$

MOMENTUM ϒ

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

NESTEROV

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

ADAGRAD

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

ADADELTA

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

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

RMSprop

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

ADAM

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

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

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

RESTRICTED BOLTZMANN MACHINES

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

where v = binary state visible

h = binary state hidden

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

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

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

$$\Delta w_{ij} = \eta . \frac{\partial}{\partial w_{ij}} \log p(v)$$
$$\Delta w_{ij} = \eta . (< v_i\ h_j >^0 - < v_i\ h_j >^1)$$

CONVOLUTIONAL NEURAL NETWORKS

$$Output\ Size = \frac{(N-F)}{S} + 1$$

where: N= input size
F = filter size
S = Stride steps
Convolution2D(N filters, filter_size, filter_size…)

SUPPORT VECTOR MACHINES

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

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

$$\lambda \to \nabla L = 0$$

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

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

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

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

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

SUPPORT VECTOR REGRESSION

$$\hat{Y} = w . < x_i \cdot x_j > + b$$

$$y - (w . < x_i \cdot x_j > + b) \leq \varepsilon$$

$$w . < x_i \cdot x_j > + b - y \leq \varepsilon$$

RIDGE REGRESSION - REGULARIZATION

$$m \coloneqq m - \frac{\sum(\hat{y} - y)^2}{N} - \frac{\lambda . m}{N}$$

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

LASSO REGRESSION - REGULARIZATION

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

$$m \to 0$$

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

SKEWNESS

Skewness < 1

KOLMOGOROV SMIRNOV

Normal sig > .005

NON PARAMETRIC

T test = Normal

Test U Mann Whitney sig < .05

CRONBACH

> .60 .70

MEDIAN

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

t TEST

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

Difference significant sig < .05

t TEST 2 SAMPLES

Levene Variância

ANOVA + 3

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

Sig < .05

TOLERANCE

Tolerance > .1

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

VARIANCE INFLATION FACTOR

VIF <10

ENTER METHOD

+ 15 cases / Variable

STEPWISE METHOD

+ 50 cases / Variable

VARIABLE SELECTION

F Test = 47 sig < .05

MISSING DATA

Delete if > 15%

DISCRIMINANT ANALYSIS

Box M sig < .05 reject H0

Wilk's Lambda sig < .05

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

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

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

ERROR MARGIN

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

ACCURACY

Confidence Interval ~ P value

HYPOTHESES TESTING

P value < .05

TRANSFORMATION OK

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

MULTICOLLINEARITY

Correlation > .90

VIF <10

Tolerance > .1

SUM OF SQUARES (explain)

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

STANDARD ERROR ESTIMATE (SEE)

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

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

MAHALANOBIS DISTANCE

same variable

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

MANHATTAN DISTANCE L

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

NET PRESENT VALUE

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

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

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

NPV=0 (IRR)

MARKOV DECISION PROCESS

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

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

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

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

ARIMA ~ NPV

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

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

ARIMA(1,1,1):

AR = number autoregressive terms

B=number non-seasonal needed for stationary

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)

and *e=noise*

PROBABILITY (coins)

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

FREQUENTIST

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

AXIOMATIC

$P(A) \geq 0$

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

PROBABILITY THEOREMS

JOIN = A or B

$P(AUB)_{EXCLUDENT} = P(A) + P(B)$

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

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

COMPLEMENTARY EVENT

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

MARGINAL PROBABILITY

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

PROBABILITY A and B

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

CONDITIONAL PROBABILITY

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

BAYES (52 cards , cancer)

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

BINOMIAL DISTRIBUTION (0,1 success)

$$P(D) = \binom{sample\ space}{sucecss}.P(s)^s.(1 - P(s))^{N-S}$$

$$P(D) = \binom{sample\ space}{sucecss}.P(s)^s.(P(\bar{s}))^{N-S}$$

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

TOTAL PROBABILITY (jars)

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

PROBABILITY k SUCCESS in n TRIALS

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

INTEGRALS

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

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

PRODUCT RULE

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

CHAIN RULE

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

INTEGRATION

$$\sum f'(x).\Delta x \begin{matrix}\Delta x = 0\\ N \to \infty\end{matrix}$$

DIFFERENTIATION

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

LINEAR ALGEBRA

ADDITION

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

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

x Vetor: Column A = Rows B

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

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

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

OR

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

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

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

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

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

IMPORTANT

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

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

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

PERMUTATION

LEFT=exchange rows

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

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

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

TRANSPOSE

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

PROPERTIES

Not commutative

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

Associative

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

Inverse (only squared)

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

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

DETERMINANT

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

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

DEMAND ELASTICITY

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