Machine Learning Mathematics
2021年1月26日 0:24

PS: Pre-Learning

<I> Linear Algebra.

- **Matrices:** Almost all Machine Learning Algorithm use Matrix Algebra in one way or another.
- **Eigenvectors and Eigenvalues:** An Eigenvector or characteristic vector of a Linear Transform T from a vector space V over a field F into itself a non-zero vector that doesn't change its direction.

  ex: $A \cdot v = \begin{bmatrix} 1 & 2 \\ 8 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \lambda \cdot v$

- **Derivates Chain Rule:**

  $F'(x) = f'(g(x)) \cdot g'(x)$ or $\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$

- **Jacobian Matrix:** The Matrix of all first-order partial derivaties of a vector-valued function.

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

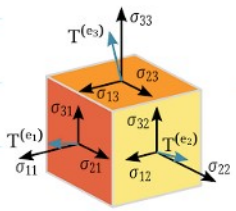
$v = J \cdot \dot{q}$

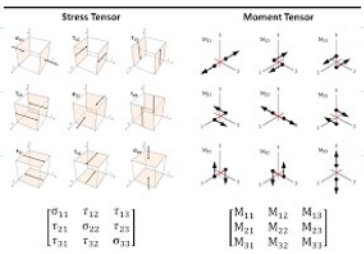
机器人关节角速度 $\dot{q}$ → 机器人末端在世界坐标系下的速度 v.

- **Gradient:** A multi-variable generalization of the derivative. The gradient is a vector-valued function.

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}$$

- **Tensors:** www.youtube.com/watch%3Fv%3Df5liqUk0ZTw





For machine Learning Purpose, a Tensor can be described as a Multimentional Matrix. Depending on the dimensions, the tensor can be a scalar, a vector, a Matrix or a Multidimentional Matrix.

Scalar (0 Tensor) $A_{\vec{0}}$, $A_{\vec{j}}$, A_k ($\vec{0}, \vec{j}, k = 0$)

Vector (1 Tensor) $A_{\vec{i}}$, $A_{\vec{j}}$, A_k ($\vec{i}, \vec{j}, k = 1$)

about n Tensor $A_{\vec{i}_1 \cdots \vec{i}_n}$ ... $A_{\vec{i}_1 \cdots}$ ... $A_{\vec{j} \cdots \vec{j}_n}$ ... $A_{k \cdots k_n}$

about n Tensor $A_{i_1 \cdots i_n} \cdots A_{\bar{i}_1 \cdots} \cdots A_{\bar{j} \cdots j_n} \cdots A_{k \cdots k_n}$

the Tensor is about the combination of n base frame Axis.

Ex: 对 $n=2$, it's Tensor is about like:

$$\text{Tensor}(2) = \begin{Bmatrix} A_{ii} & A_{ij} & A_{ik} \\ A_{\bar{j}i} & A_{\bar{j}j} & A_{\bar{j}k} \\ A_{ki} & A_{kj} & A_{kk} \end{Bmatrix}$$

and A_{ii} is $\vec{i} \to \vec{i}$ like this.

A_{ij} is $\vec{i} \to \vec{j}$

(II) STATISTICS.

- Mean (平均数) → the average of the Inputs.
- Median (中位数) → Value in the middle or an ordered list. or an average of two in middle.
- Mode (众数) → Most Freqyency existed value. qually.
- Quantile (分位数) → Dividing observation numbers in a sample list

→ Dispersion (分散)

- Range (范围)
- Medium Absolute Deviation (与中间值的绝对值散差)