Machine learning with Math

Useop Gim

January 8, 2023

Chapter 1

An introducion to Linear algebra

1.1 Basic

Definition 1. *Group*
Closure, associativity, identity, inverse
If commutativity, abelian

Definition 2. *Ring*
operation(+,$\times$)
Closure,associativity, identity(+), inverse(+), distributivity
If commutativity, commutativity ring

Definition 3. *Filed*
operation + and times
Closure, commutativity, associativity, identity, inverse, distributivity

Definition 4. *Vector Space*
operation + and scalar times
Closure, commutativity, associativity, identity, inverse, distributivity

Definition 5. *inner product*
Linearity
Conjugate symmetry
Positive-definiteness

Definition 6. *Norm*
Triangle inequality
Absolute homogeneity
Positive-definiteness
All norm is seminorm but not all seminorm(nonnegativity) is Norm

Definition 7. *Transformation*
Let V and V' are vector sapce

$$T : V \to V'$$

Linear Transformation is specific transformation which transforms to linear equation

Definition 8. *Matrix*
Matrix is the system for linear polynomial equation in vector space

1.2 Basic for Vector

Definition 9. *Elementary row operation matrix*
1. Bilinear and Scalar multiplication $R_i(x) = xR_i$
2. Swap $R_ij \Leftrightarrow R_i \leftrightarrow R_j$
3. Addition $R_ij(x) = R_i + xR_j$

Theorem 1. *Row Echelon form elimination and Gaussian Jordan elimination*
Both elimination augmentes the matrix
For RREF, the first non-zero row value is "leading variable"
The zero value for the row or without introduced variable is "free variable"

Theorem 2. *linear equation system of vector space*
1. No solution
2. One solution
3. Infinity many solution

Definition 10. *Homogeneous equation*
It is an equation to zero, whose each term contains the function or 'one of its derivatives'.

Definition 11. *Linearly independent*
If the solution of homogeneous equation has only zero for all terms, it is LD
And we call the solution is trivial solution

Definition 12. *Invertible in Matrix*
1. Linear transformation of A is bijective
2. Since LT of A is bijective, RREF is Identity
3. Since RREF no free variable, Elementary row operation matrix times A is identity
4. Since no free variable, $Ax = 0$ has only trivial solution$\Leftrightarrow$ Linearly independent
5. Since no free variable, $Ax = b$ has only one solution
6. $rank(A) = n$
7. $Null(A) = 0$
8. $def(A) \neq 0$
9. $0 \notin eigen(A)$
10. A is non singular

Definition 13. *Transpose*

$$[A^T]_{ij} = A_{ji}$$

Definition 14. *determinant*
Let Mutually Exclusive Collectively Exhaustive by permutation

$$\sum_{i=1} \left(sign(\epsilon_i) \prod_{j=1} [A]_{j\epsilon_i} \right)$$

Definition 15. *Cofactor expansion(Laplace expansion) and Adjugate of matrix*

$$adj(A) = C^T$$

Definition 16. *Orthogonal*
Inner product of two different vector gives zero

Definition 17. *Funder mental subspace*
Subspce : subset for vector space with linearity
Span : the set of vectors which constructs subspace(S) from linear equation system
Basis : the set of vectors in linearly independent span(S)
Dim : the maximal order of linear independent subset $dim(S) = |basis(S)|$
Row(S) : basis for row space of S thus $row(S) = col(S^T)$
Col(S) : basis for col space of S thus $col(S) = row(S^T)$
Null(S) : null space is basis of linearly independent equation
In RREF : row(S) is the leading variables and null is the free variables
Null(S^T*) : WLG same as above*
Nullity(S) : dim(null(A))
rank(S) : The largest dimension of basis of row space $rank(S) = dim(row(S))$
usually rank(S) from row space
Ker(A) : $x : G(x) = e_y$
In matrix linear system $Ker(A) \leftrightarrow Ax = 0$
Therefore $Ker(A) = Null(A)$
Img(A) : $y : G(x) = y$ *Therefore* $Img(A) = col(A)$

Theorem 3. *The relationship between subspace*
$row(S) = col(S^T)$
$row(S) \perp Null(S^T)$
$col(S) = row(S^T)$
$col(S) \perp Null(S^T)$

Theorem 4. *Rank and Nullity theorem*
We know $dim(Img(A)) + dim(Ker(A)) = dim(domain(A))$
Since 1st isomorphism theorem from $V/Ker(A)$ *to* $Img(A)$ *(bij,homo)*

$$rank(A) + Nullity(A) = n$$

Theorem 5. *rank*
We know that $rank(A) + Nullity(A) = n$
And $rank(A) = dim(row(A))$
Also, $Ker(A) = Null(A)$ *and* $Img(A) = col(A)$
$dim(Img(A)) + dim(Ker(A)) = dim(domain(A))$
$\Leftrightarrow dim(Img(A)) + dim(Ker(A)) = n$
$\Leftrightarrow dim(Img(A)) + dim(Null(A)) = n$
$\Leftrightarrow dim(col(A)) + Nullity(A) = n$
$\Leftrightarrow dim(col(A)) + n - rank(A) = n$
$\Leftrightarrow dim(col(A)) = rank(A)$
$\Leftrightarrow dim(col(A)) = dim(row(A))$
it doesn't mean col(A) = row(A) it just indicates the rank (dim)

Definition 18. *Eigen value*
eigen value is obtained from the solution $Ax = \lambda x \to \lambda I - A = 0, x \neq 0$

Definition 19. *Characteristic equation*

$$\lambda I - A = 0$$

Definition 20. *Characteristic polynomial equation*

$$P_A(\lambda) = det(\lambda I - A)$$

Definition 21. *Eigen vector*
The basis of $(\lambda I - A)x = 0$ *with* $x \neq 0$

Theorem 6. *Eigen space is Singular*
If eigenvector is existed, then

$$det(\lambda I - A) = 0$$

If $P_A(\lambda) = 0$ *then* $(\lambda I - A)x = 0 \Rightarrow x = (\lambda I - A)0$
However, $x \neq 0$ *and so it is contradiction*

Theorem 7. *Zero Eigen value*
If eigen value is zero then $(\lambda I - A)x = 0 \Rightarrow Ax = 0$
Thus it is linearly independent ($x \neq 0$*)*
Hence x be the null of(A)

$$\lambda = 0$$

$$\Rightarrow Ax = 0 \Rightarrow x \in Null(A) \wedge x \neq 0$$

$$\Rightarrow Null(A) \text{ has no trivial solution}$$

$$\Rightarrow rank(A) + Nullity(A) = n \to rank(A) \in [0, n)$$

Definition 22. *Nilpotent*
The nilpotent matrices satisfies $M^n_{m\times m} = 0$ *for* $\exists n \in \mathbb{N}$

$$Ax = \lambda x \Rightarrow A^2x = A(Ax) = A(\lambda x) = \lambda^2 x$$

Through induction

$$A^n x = \lambda^n x \Rightarrow (A^n - \lambda)x = 0$$

We know that $A^n = 0$
Thus $\lambda = 0$ *Therefore all eigenvalues(*λ*) of it are 0*
Hence $rank(Nilpotent) \in [0, m)$

Definition 23. *Trace*
Trace is the sum of main diagonal(only sauqre) elements

Theorem 8. *Tr(A) and det(A)*
Tr(A) $= \sum_i^n \lambda_i$
Because Characteristic equation is 0

$$tr(\lambda I - A) = \sum_i (\lambda_i - a_{ii})$$

$$0 = \sum_i \lambda_i - tr(A)$$

$$tr(A) = \sum_i \lambda_i$$

det(A) $= \prod_i^n \lambda_i$ *Because Characteristic polynomial equation is zero (non-invertible)*

$$P_A(\lambda) = det(\lambda I - A)$$

Since λ *is root of polynomial* P

$$P_A(\lambda) = \prod_i (\lambda - \lambda_i)$$

Through above two equation

$$det(\lambda I - A) = \prod_i (\lambda - \lambda_i)$$

Then let $\lambda = 0$

$$det(A) = \prod_i (\lambda_i)$$

Definition 24. *Algebraic multiplicity and Geometric multiplicity*
$am(\lambda)$ *is the number of eigenvalue*
$gm(\lambda)$ *is the number of eigenvector*

1.3 Applied Decomposition

Chapter 2

An introduction to ML

Theorem 9. *Topological meaning of the neural network*
Vector space, Hyperplane, optimization, statical

Chapter 3

Introduction of Probability and Statics

3.1 Basic

Definition 25. *Multi-variable function*

$$f : \mathbb{C}^n \to C$$

Definition 26. *Probability Density function*
When $F_x(x) = \int_a^b p(x)dx$

$$P(x) = \frac{dF_x(x)}{dx}$$

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(X \leq b) - F_X(X \leq a) = \int_a^b p_X(x)dx$$

it is continuous
It follows Lebesque integral

Definition 27. *Probability Mass function*

$$P(a \leq X \geq b) = \sum_{x_i} f(x_i)$$

it is discrete

Definition 28. *Cumulative Distribution function*

$$F_X(x) = P\{X \leq x\}$$

There are two type of cdf, discrete$\sum$ *and continuous*$\int$

Definition 29. *Joint Cumulative Distribution function*

$$F_{XY}(x,y) = P(X \leq x) \cap (Y \leq y)$$

Thus for PDF,

$$F_{XY}(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} p_{XY}(u,v)dudv = \int_{-\infty}^{y}\int_{-\infty}^{x} p_{XY}(u,v)dvdd$$

Definition 30. *Independence*

$$P(A,B) = P(A)P(B)$$

Definition 31. *Conditional Probability*

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

Definition 32. *Bayesian Probability*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

In pdf with joint probability

$$p_{Y|X}(y|x) = \frac{p_{X|Y}(x|y)p_Y(y)}{\int_{-\infty}^{\infty} p_{X|Y}(x|y)p_Y(y)dy}$$

Definition 33. *Marginal density function*
From the JCDF, we can obtained the pdf for X and Y from below

$$p_x(x) = \int_{-\infty}^{\infty} p_{XY}(x,y)dy$$

$$p_Y(y) = \int_{-\infty}^{\infty} p_{XY}(x,y)dx$$

Definition 34. *Independent and identically distributed*
IID means each sample is independent with uniformly picked

Definition 35. *Dirac delta function*

$$\int_{-\infty}^{\infty} \delta(x)f(x)dx = f(0)$$

$$\delta(x) = \begin{cases} \infty(x=0) \\ 0(x \neq 0) \end{cases}$$

3.2 Statistics

Definition 36. *Expectation*

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x p_X(x) dx$$

For random variable

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} g(x) p_X(x) dx$$

For joint

$$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) p_X(x,y) dx dy$$

For random vector

$$\mathbb{E}[X] = \begin{bmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_n] \end{bmatrix} = \int_{-\infty}^{\infty} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} p_X(x) dx$$

For conditional expectation, let $x \in X$

$$\mathbb{E}[X|Y] = \int_{-\infty}^{\infty} x p_{x|Y}(x|Y) dx$$

Definition 37. *Variance*

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

If X and Y are independent

$$\begin{aligned} Var(X+Y) &= \mathbb{E}[(X + Y - (\mu_X + \mu_Y)^2] \\ &= \mathbb{E}[((X - \mu_X) + (Y - \mu_Y))^2] \\ &= \mathbb{E}[(X - \mu_X)^2 + (Y - \mu_Y)^2 + 2(X - \mu_X)(Y - \mu_Y)] \\ &= \mathbb{E}[(X - \mu_x)^2] + \mathbb{E}[(Y - \mu_y)^2] + 2\mathbb{E}[(X - \mu_x)(Y - \mu_y)] \\ &= Var(X) + Var(Y) + 2\mathbb{E}[(X - \mu_x)(Y - \mu_Y)] \end{aligned}$$

Definition 38. *Covariance*

$$Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

For random vector

$$Cov(X) = \mathbb{E}[(X - \mathbb{E}[x])(X - \mathbb{E}[X])^T] = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])(x - \mathbb{E}[X])^T p_x(x) dx$$

It represented as matrix which is symmetric matrix ($A^T = A$)
For conditional covariance for the random vector

$$Cov(X|Y) = \mathbb{E}[(X - \mathbb{E}[X|Y])(X - \mathbb{E}[X|Y])^T|Y]$$

Definition 39. *Correlation*

$$Cor(X,Y) = \mathbb{E}[XY] = \int_{-\infty}^{\infty} xyp_{XY}(x,y)dxdy$$

Definition 40. *Correlation Coefficient*

$$p_{XY} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

It is measurement of how likehood $(X - \mathbb{E}[X])$ *and* $(Y - \mathbb{E}[Y])$

Definition 41. *Random vector marginal density function*

$$F_x(x) = \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_n} p_{X_1...X_n}(u_1, ..., u_n)du_1...du_n$$

Short form, let $u = [u_1...u_n]^T$

$$F_x(x) = \int_{-\infty}^{X} p_X(u)du$$

Definition 42. *Random vector conditional probability density function*

$$P(X|Y) = \int_{-\inf}^{x} p_{X|Y}(u|y)du$$

Definition 43. *Population and Sample*
The variance of the Population and Sample

$$\sigma^2 = \frac{\sum(X_i - \mu)^2}{N}$$

$$s^2 = \frac{\sum(X'_i - \mu')^2}{n-1} \ \forall X' \in X$$

Definition 44. *The average of Sample*
If we obtained the sample through IID (Independent and identically distributed)

$$\begin{aligned}
\mathbb{E}[f(X)] &= \int_{-\infty}^{\infty} f(x)p_X(x)dx \\
&= \int_{-\infty}^{\infty} f(x) \sum_{i=1}^{N} \frac{1}{N} \delta(x - x^{(i)})dx \ \textit{Through dirac delta} \\
&= \frac{1}{N} \sum_{i=1}^{N} f(x^{(i)})
\end{aligned}$$

Then if we let the $f(X) = X$ *for random variable* X
The expected μ_X

$$\mu_X = \mathbb{E}[X] \approx \frac{1}{N} \sum_{i=1}^{N} x^{(i)}$$

Definition 45. *The mean of samples equals to the mean of population*

$$\mathbb{E}(\bar{X}) = \mu$$

Let $\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$ *and* $x \in X$

$$\begin{aligned}
\mathbb{E}(\bar{X}) &= \mathbb{E}\left(\frac{\sum_{i=1}^{n} x_i}{n}\right) \\
&= \frac{1}{n}\mathbb{E}\left(\sum + i = 1^n x_i\right) \\
&= \frac{1}{n}\mathbb{E}\left(x_1 + x_2 + ... + x_n\right) \\
&= \frac{1}{n}\left(\mathbb{E}(x_1) + \mathbb{E}(x_2) + ... + \mathbb{E}(x_n)\right) \\
&= \frac{1}{n}(n\mu) \\
&= \mu
\end{aligned}$$

Definition 46. *The variance of the samples equals to the* $\frac{1}{N}Var[X]$

$$\begin{aligned}
Var[\bar{X}] &= \mathbb{E}[(\bar{X} - \mathbb{E}[\bar{X}])^2] \\
&= \mathbb{E}[(\bar{X} - \mu)^2] \\
&= \mathbb{E}\left[\left(\frac{1}{N}\left(\sum_{i=1}^{N} X_i - N_\mu\right)\right)^2\right] \\
&= \mathbb{E}\left[\frac{1}{N^2}\sum_{i=1}^{N}\sum_{i=1}^{N}(X_i - \mu)(X_j - \mu)\right] \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}[(X_i - \mu)(X_j - \mu)]
\end{aligned}$$

Then the sample of i *and* j *are independent*

$$E[(X_i - \mu)(X_j - \mu)] = 0(i \neq j)$$

Therefore

$$\begin{aligned}
Var[\bar{X}] &= \frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}[(X_i - \mu)^2] \\
&= \frac{1}{N^2}N\mathbb{E}[(X - \mu)^2] \\
&= \frac{1}{N}\mathbb{E}[(X - \mu)^2] \\
&= \frac{1}{N}Var[X]
\end{aligned}$$

Definition 47. *The value of variance of expected samples*

$$\mathbb{E}[S^2] = \frac{N-1}{N}\sigma^2$$

Thus the variance of expected samples is less then variance of population

Definition 48. *Skewness*
The measurement of asymmetry

$$\mathbb{E}\left[\left(\frac{X-\mu}{\gamma}\right)^3\right] = \frac{\mu^3}{\gamma^3}$$

Definition 49. *Kurtosis*
The measurement of tailedness of distribution(from the center)

$$\mathbb{E}\left[\left(\frac{X-\mu}{\gamma}\right)^4\right] = \frac{\mu^4}{\gamma^4}$$

Definition 50. *Moment*
The characteristic of expectation, variance and skewness If moments are the same they are the same distribution

$$\mu = \mathbb{E}[(X-\mu)^n] = \int (x-\mu)^n p(x)dx$$

Chapter 4

Probability and statical model

Definition 51. *Bernoulli, Binomial distribution*

$$Ber(\theta)$$

$$f(x_i|\theta) = \begin{cases} \theta(x_i = 1) \\ 1 - \theta(x_i = 0) \end{cases}$$

$$E(X_i) = \theta, Var(X_i) = \theta(1-\theta)$$

Definition 52. *Bernoulli, Binomial distribution*

$$Ber(\theta)$$

$$f(x_i|\theta) = \begin{cases} \theta(x_i = 1) \\ 1 - \theta(x_i = 0) \end{cases}$$

$$f(x_i|\theta) = \prod_{i=1}^{n} \theta_i^{x_i}(1-\theta_i)^{1-x_i}$$

When it measures the n times success do

$$Bin(m, \theta)$$

$$f(y|\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y}, y \in \mathbb{Z}, \theta > 0$$

Definition 53. *Possion distribution*

$$P(X = x) = f(x|\lambda) = e^{-\lambda}\frac{\lambda^x}{x!}, x = 0, 1, 2, ..\lambda > 0$$

$$\mathbb{E}(X) = \lambda, Var(X) = \lambda$$

Theorem 10. *When* $n \to \infty \wedge p \to 0 (np \to \lambda)$ *Then* $Bin(n, \theta) = P_0(\lambda)$

Definition 54. *Geometric distribution*

$$Gem(X)$$
$$P(X = x) = (1-p)^x p, x \in \mathbb{Z}$$
$$\mathbb{E}(X) = \frac{1-p}{p}, Var(X) = \frac{1-p}{p^2}$$

Definition 55. *Negative-Binomial distribution*

$$P(X = x) = \genfrac{}{}{0pt}{}{x+k-1}{k-1} p^k (1-p)^x, x = 0, 1, 2, ...$$
$$\mathbb{E}(X) = \frac{k(1-p)}{p}, var(X) = \frac{k(1-p)}{p^2}$$

Definition 56. *Multinomial distribution*
When picked ball at i *which color is* j *and the number of color* j *is* n_j

$$Mul(p_1, ..., p_k)$$
$$P(X_i = n_i, ..., X_k = n_k) = \frac{n!}{n_1! n_2! ... n_k!} \prod_{j=1}^{k} p_j^{n_j}, \sum n_j = n$$

Definition 57. *Uniform distribution*

$$f(x|\theta_1, \theta_2) = \begin{array}{c} \frac{1}{\theta_2 = \theta_1}, \theta_1 < x < \theta_2 \\ 0 \end{array}$$
$$\mathbb{E}(X) = \frac{\theta_1 + \theta_2}{2}, Var(X) = \frac{(\theta_2 - \theta_1)^2}{12}$$

Definition 58. *Exponential distribution*

$$Exp(\theta)$$
$$f(x|\theta) = \theta e^{-\theta x}, x > 0, \theta > 0$$
$$\mathbb{E}(X) = \frac{1}{\theta}, Var(X) = \frac{1}{\theta^2}$$

Definition 59. *Gamma distribution*

$$Gam(\alpha, \beta)$$
$$f(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha - 1} e^{-\beta x}, x > 0, \alpha, \beta > 0$$
$$\mathbb{E}(X) = \frac{\alpha}{\beta}, Var(X) = \frac{\alpha}{\beta^2}$$

If X *are independent and* $Exp(\beta)$, $Y = \sum X_i$
The Y *gives*

$$Gam(n, \beta)$$

Definition 60. *Chi-Square distribution*
When Gamma distribution has $\alpha = \frac{v}{2}$ *and* $\beta = \frac{1}{2}$, *this has freedom v chi-square* X_v^2 *distribution*

$$\mathbb{E}(X) = v, Var(X) = 2v$$

Definition 61. *Inverse Gamma distribution*
When $Y = Gam(\alpha, \beta), Z = frac1Y$
Through Newton-Raphson

$$f(z) = \frac{\beta}{\Gamma(\alpha)} z^{-(\alpha+1)} e^{\frac{-\beta}{z}}$$

$$\mathbb{E}(Z) = \frac{\beta}{(\alpha - 1)}, Var(Z) = \frac{\beta}{(\alpha - 1)^2(\alpha - 2)}, \alpha > 2$$

Definition 62. *Normal distribution*

$$f(x|\theta) = \prod \frac{1}{\sqrt{2\pi}\gamma} e^{\frac{(x_i - \mu)^2}{2\gamma^2}}$$

Chapter 5

An introduction to RL

5.1 Basic

Theorem 11. *Markov Property*
It is the probability of event which is related with the past events.

$$P(x_t|x_0, x_1, ..x_{t-1}) = P(x_t)$$

The Markov model is based on the above property.

Theorem 12. *Markov Process(Markove Chain)*
In the Markov process is a process which following the Markov Property with tuple (S, P)

Theorem 13. *State transition Probability* P
The probability making transform state is called "state transition probability."

Theorem 14. *Reward* R
The reward in Markov Process is obtained from transition.

Theorem 15. *Markov Decision Process*
The Markov Decision Process represent the Markov Process as tuple.

Theorem 16. *Markov Reward Process*
It adds a reward and a discount factor element in process (S, P, R, γ)

Theorem 17. *Markov Decision Process*
It adds an action element in process (S, A, P, R, γ)

Theorem 18. *Policy* π
From the Markov Process, policy is the distribution of all action by current state.

$$\pi(a|s) = P\left[A_t = a|S_t\right]$$

Theorem 19. *State-value*
Let G is total future reward
The state-value is the value of current state

$$v(s) = \mathbb{E}\left[G_t | S_t = s\right]$$

Theorem 20. *Action-value*
The action-value is the value of action for current state

$$q(s,a) = \mathbb{E}\left[G_t | S_t = s, A_t = a\right]$$

Theorem 21. *Bellman Expected equation*

$$\begin{aligned} v_\pi(s) &= \mathbb{E}\left[R_{t+1} + \gamma v(s_{t+1}) | s_t\right] \\ q_\pi(s,a) &= \mathbb{E}\left[R_{t+1} + \gamma q(s_{t+1}, a_{t+1}) | s_t, a_t\right] \end{aligned}$$

Theorem 22. *Bellman Optimality equation*

$$\begin{aligned} v_*(s) &= max\mathbb{E}\left[R_{t+1} + \gamma v(s_{t+1}) | s_t\right] \\ q_*(s,a) &= max\mathbb{E}\left[R_{t+1} + \gamma q(s_{t+1}, a_{t+1}) | s_t, a_t\right] \end{aligned}$$

Theorem 23. *Transition by policy*
The probability of transition from state s_t to s_{t+1} with policy

$$P^\pi(s_{t+1}|s_t) = \sum_{a \in A_t} \pi(a|s_t) p(s_{t+1}|s_t, a) r(s,a)$$

Theorem 24. *Reward by policy*
The probability of transition from state s_t to s_{t+1} with policy

$$R^\pi(s_t, a) = \sum_{a \in A_t} \pi(a|s_t) r(s_t, a)$$

5.2 Policy Gradient Theorem

Theorem 25. *State-value by policy*
First convert State-value function as policy form

$$\begin{aligned} v_\pi(s_t) &= \mathbb{E}_\pi\left[R_{t+1} + \gamma v(s_{t+1})\right] \\ &= R^\pi_{s_t} + \gamma \sum_{s_t} v(s_{t+1}) \\ &= \sum_{a \in A} \pi(a|s_t) r(s_t, a) + \gamma \sum_{s_{t+1} \in S} \sum_{a \in A_t} \pi(a|s_t) p(s_{t+1}|s_t, a) v_\pi(s_{t+1}) \end{aligned}$$

Through the summation property

$$v_\pi(s_t) = \sum_{a \in A_t} \pi(a|s_t) r(s_t, a) + \gamma \sum_{a \in A_t} \pi(a|s_t) \sum p(s_{t+1}|s_t, a) v_\pi(s_{t+1})$$

Theorem 26. *Action-value by policy*

$$\begin{aligned} q_\pi(s_t, a) &= \mathbb{E}\left[R_{t+1} + \gamma q(s_{t+1}, a_{t+1}) | s_t, a_t\right] \\ &= r(s, a) + \gamma \sum_{s_{t+1} \in S} p(s_{t+1} | s, a) \sum_{a' \in A} \pi(a' | s_{t+1}) q_\pi(s_{t+1}, a') \end{aligned}$$

Theorem 27. *State-value and Action-value relationship*

$$v_\pi(s_t) = \sum_{a_t \in A_t} \pi(a_t | s_t) q_\pi(s_t, a_t)$$

$$q_\pi(s_t, a) = r(s_t, a) + v_\pi(s_t)$$

Theorem 28. *Policy Gradient Theorem*
For the stochastic policy π

$$\begin{aligned} \nabla v_\pi &= \nabla \left(\sum_a \pi q \right) \\ &= \sum_a \left(\nabla \pi(a|s) q_\pi + \sum \pi(a|s) \nabla q_\pi \right) \\ &= \sum_a \left(\nabla \pi(a|s) q_\pi + \pi(a|s) \sum_{s'} \nabla \left(\sum_{s'} P^a_{s's} \cdot q_\pi(s', a') \right) \right) \\ &= \sum_a \left(\nabla \pi(a|s) q_\pi + \pi(a|s) \sum_{s'} \nabla \left(\sum_{s'} P^a_{s's} \cdot (r + v_\pi(s')) \right) \right) \\ &= \sum_a \left(\nabla \pi(a|s) q_\pi + \pi(a|s) \sum_{s'} P^a_{s's} \cdot \nabla v_\pi(s') \right) \end{aligned}$$

Therefore

$$v_\pi(s) = \sum_a \left(\nabla \pi(a|s) q_\pi(s, a) + \pi(a|s) \sum_{s'} P^a_{s's} \cdot \nabla v_\pi(s') \right)$$

Then the $v_\pi(s,a)$ is repeated Thus let $\phi(s) = \sum_a \nabla\pi(a|s)q_\pi(s,a)$

$$
\begin{aligned}
\nabla v_\pi(s) &= \phi(s) + \sum_a \left(\pi(a|s) \sum_{s'} P^a_{s's} \cdot \nabla v_\pi(s') \right) \\
&= \phi(s) + \sum_a \sum_{s'} \pi(a|s) P^a_{s's} \cdot \nabla v_\pi(s') \\
&= \phi(s) + \sum_{s'} \sum_a \pi(a|s) P^a_{s's} \cdot \nabla v_\pi(s') \\
&= \phi(s) + \sum_{s'} \sum_k p_\pi(s \to s', k) \cdot \nabla v_\pi(s') \\
&= \phi(s) + \sum_{s'} \sum_k p_\pi(s \to s', k) \cdot (\phi(s') + ...) \\
&= \phi(s) + \sum_x \sum_{k=0} p_\pi(s \to x, k) \cdot \phi(x)
\end{aligned}
$$

Andso let $\sum_{k=0} p_\pi(s \to x, k)$ as $\eta(s)$

$$
\begin{aligned}
\nabla v_\pi &= \sum_s \eta(s)\phi(s) \\
&= \left(\sum_s \eta(s) \right) \sum \frac{\eta(s)}{\sum \eta(s)} \phi(s) \\
&= \sum_s d_\pi(s)\phi(s) && \textit{Since } \textstyle\sum \eta \textit{ is constant} \quad \textit{d is the stationary distribution} \\
&= \sum_s d_\pi(s) \sum_a \nabla\pi(a|s)q_\pi(s,a)
\end{aligned}
$$

Hence the derivative of the expected value function is obateind from the gradient of policy without taking derivative for the reward function

Chapter 6

An introducion to Data

6.1 Basic

Definition 63. *DP, MC, TD*
DP is dynamic programming, it uses the model base
MC is Monte Carlo, it uses bunch of samples then estimates the probability (ex calculate circle of pi)
TD is Temporal Difference, it uses the difference of the transitional value from one step behind state

Definition 64. *Exploitation vs Exploration*
Exploitation is deciding the best action through the using the given samples
Exploration is collecting samples

Definition 65. *TP, TN, FP, FN, SE, SP, FPR, ROC, AUC*
TP is true positive
TN is true negative
FP is false postive (type 1 error)
FN is false negative (type 2 error)
SE is Sensitivity $\frac{TP}{TP+FN}$ *which is the rate of correct positive*
SP is Specificity $\frac{TN}{TN+FP}$ *which is the rate of correct negative*
FPR is False Positive Rate which is the rate of wrong positive
ROC is the the curve by vertical SE horizontal FPR or (1-SP)
AUC is the area of under ROC