eigen of R_x is eigen of R_Y

$$R_Y = \hat{R}_X + \beta I$$

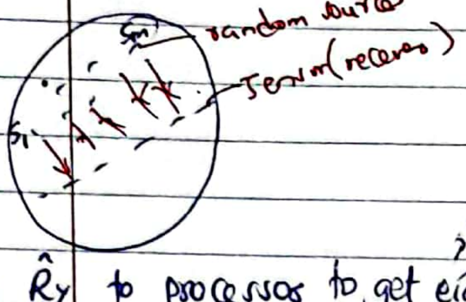
Implementation of Algo

- observe samples from $Y : Y^{(1)} \dots Y^{M}$. $M \gg n$

* This is learning phase.

- estimate R_Y → covariance matrix of received signal

$$R_Y = \mathbb{E}[\underline{Y} \cdot Y^T] \simeq \frac{1}{M} \sum_{k=1}^{M} Y^{(k)}_{,} \cdot Y^{(k)T} = \hat{R}_Y$$



- Give $\hat{R}_Y$ to processor to get eigen vectors

→ orthogonal (symmetric)

$\hat{R}_y \rightarrow \underline{u}_i, \underline{u}_2 \dots u_m, u_{m+1} \dots u_n$ eigen vectors are arranged in decreasing order

$\rightarrow \lambda_1 \geq \lambda_2 \geq \dots \lambda_m \geq \lambda_{m+1} =$ positive always

if

$$R = \mathbb{E}[zz^T]$$

$$\underline{u}^T R u = u^T \mathbb{E}[zz^T] u = E[\underbrace{u^T \cdot zz^T u}_{}] \geq 0$$

↓ scalar quantity which is always positive

$$Rv = \lambda v$$

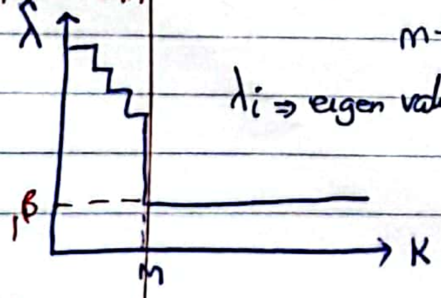
$$v^T R v = \lambda v^T v = \underline{\lambda \cdot vv^2} \geq 0$$

hence

$$\underbrace{\lambda_1 \geq \lambda_2 \geq \lambda_m}_{\text{span } S_m} \dots \geq \underbrace{\lambda_{m+1} \dots \lambda_N}_{\beta}.$$

→ dominant eigen vector

β → value of minimum eigen value



m → No of source

λ_i ⇒ eigen values of $R_Y = YY^T$ in decreasing order

$m \gg N$.

Covariance of Matrix.

$$L = \sum_{i=1}^{M} \left(1 - \frac{\hat{\beta}}{\lambda_i}\right) \hat{\underline{u}}_i \underline{u}_i^T$$ → like Band Pass bandpass filter.

---

error

$\varepsilon = X - \hat{X} \Rightarrow \operatorname{Var} \varepsilon = \operatorname{tr} E[\varepsilon \varepsilon^T]$.

since $Y = X + W$ → signal & noise uncorrelated.

$$\operatorname{Var} \varepsilon = \operatorname{tr}[R_x - R_x \cdot R_Y^{-1} R_x].$$

$$= \sum_{i=1}^{m} \left[\lambda_i - \frac{\lambda_i^2}{\lambda_i + \beta}\right].$$

$$= \beta \sum_{i=1}^{M} \frac{\lambda_i}{\lambda_i + \beta} = \beta \sum_{i=1}^{M} \frac{\lambda_i / \beta}{\frac{\lambda_i}{\beta} + 1} \leq m\beta$$

→ SNR

Note

$\operatorname{tr}(ww^T) = N\beta$

Blue - original

Red - estimated signal

plot (u(:,1)) → eigen vector
(u(:,2)) (sinusoid).

plot (u(:,3)) → obtain noise.

→ spectral density of stationary signal.

Application : Maximum Entropy Estimation of Spectrum.

→ If you have any stationary 2^{nd} order process $\{X_1, X_2 ..\}$

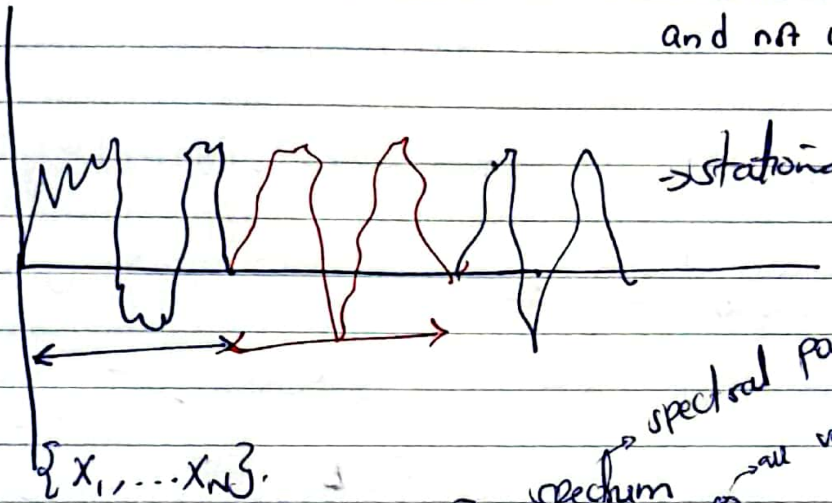
$EX_i = 0$ → zero mean ; $E|X_1|^2 < \infty$ → finite variances.

$$E[X_i X_{i+m}] = \overset{\text{correlation}}{C_n}. \quad |C_n| < \infty \text{ is finite if the variance is finite}$$

$$\underbrace{E[X_i \cdot X_j]^2}_{<\infty} \leq \underbrace{E[X_i^2]\, E[X_j^2]}_{<\infty}.$$

finite

④ ears are sensitive to spectrum and not waveform



→ stationary e.g. speech

$\{X_1, ... X_N\}$.

spectral power density

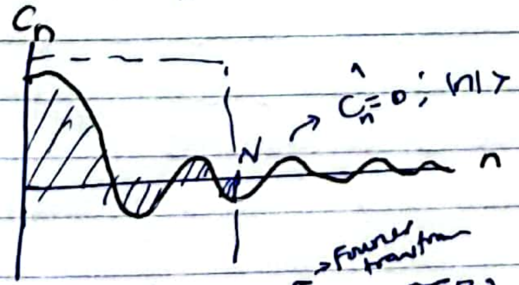
Reminder: If $C_n = E[X_K \cdot X_{K+n}]$ → $\overset{\text{spectrum}}{\Gamma(\lambda)} = \sum_{n=-\infty}^{\infty} C_n e^{-i2\pi\lambda n}$ → all values of C_n

gamma

$$\hat{\Gamma}(\lambda) = f(X_1, X_N) \rightarrow \text{estimate of spectrum}$$

$$\hat{\Gamma}(\lambda) = \sum_{n=\infty}^{\infty} \hat{C}_n \cdot e^{-i2\pi\lambda n}$$

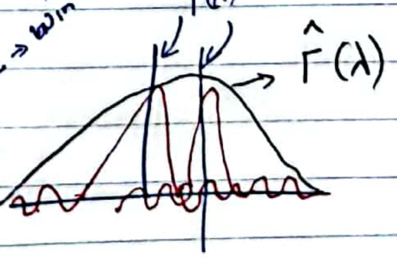
$$\hat{C}_n = \frac{1}{n-1} \sum_{K=1}^{n-1} X_K X_{K+n} \rightarrow \text{Arithmetic mean.}$$



→ $\hat{C}_n = 0 ; |n| > N$

$n > N$ → we can estimate $\overset{\text{portion}}{\wedge}$ up to N

→ sinc → window



$$\hat{C}_n = C_n \cdot \underset{\substack{\downarrow \\ \text{Rectangular window}}}{\Phi_n} \xrightarrow[\;]{F} \Gamma[\lambda] * \tilde{\Phi}(\lambda)$$

→ Fourier transform

$$0 \leq E[(X-\alpha Y)^2] = EX^2 - 2\alpha E[XY] + \alpha^2 E[Y^2]$$
$$= C - 2\alpha B + \alpha^2 A$$
$$\rightarrow B^2 - AC \geq 0$$

Maximum entropy, $\rightarrow$ finite process

$\{X_i\}_{i=1}^{\infty} \Rightarrow$ stochastic process

$f(X_{1:n}) = \exp \frac{1}{2} x^T C^{-1} x$

$C_{ij} = E[X_i X_j]$

$f(X_{1:n}]$, $n < 1, 2 \ldots \rightarrow$ Partial distributions

Differential entropy of source; $\rightarrow$ continuous process

$$h_n(X_1 \ldots X_n) = h_n(X_{1:n}) = -\int f(X_{1:n}) \log f(X_{1:n}) dx^n = h(f) \quad \text{(i)}$$

$$h = \lim_{n\to\infty} \frac{h_n(X_{1:n})}{n} \rightarrow \text{Exists for stationary source.} \quad \text{(ii)}$$

*For Gaussian process $h = \frac{1}{2}\log(2\pi\Gamma] + \frac{1}{4\pi}\int_{-\pi}^{\pi} \log[2\pi\Gamma(\lambda)] d\lambda$

$$\Gamma(\lambda) = \frac{1}{2\pi} \sum_{\lambda=-\infty}^{\infty} \sigma(\lambda) e^{-i\lambda l} = \text{spectral density} \quad \text{(iii)}$$

$$\sigma(l) = E[X_n X_{n\pm l}] = C_l \quad \text{---- (iv)}$$

Problem: Estimate $\Gamma(\lambda)$ or $\sigma(l)$ for all $L = 0, 1, 2 \ldots$ given subject to constraints. i.e. $\begin{cases} \sigma(0) = \alpha_0, \ldots \sigma_{(p)} = \alpha_p, \hat{\sigma}(p+1), \hat{\sigma}(p+2) \ldots \\ \{\sigma(L)\}_{L=\infty}^{\infty} \text{ Auto covariance} \end{cases}$

Theorem: The stochastic process $\{X_i\}$ that maximizes the differential entropy subject to correlation constraint $E[X_i X_{i+K}] = \alpha_K$; -(V) $K = 0:p$ & $i = 1, 2$ ..time is p^{th} order gauss Markov Process satisfying these constraints.

Rem: $\{X_i\}$ is not assumed to be gaussian, non-stationary.

Proof

Suppose $X_{1:n}$ is any Collection of r.v.s satisfying (v) and Consider an auxellary process $Z_{1:n}$ with Zero mean, Normal with covariance given by(σ)

$Z'_{1:n}$ is p^{th} order gauss-markov process satisfying (v)

$$Z'_K = g_1 Z'_{K-1} + \cdots + g_p Z'_{K-p} + \overset{\to \text{iid}}{U_K}$$

$g \to$ are fixed coefficients.

then for $n \geq p$

$$h(X_{1:n}) \overset{(a)}{\leq} h(Z_{1:n})$$

$$\leq h(Z_{1:p}) + \sum_{K=p+1}^{n} h(Z_K | Z_{1:K-1}) \quad (b)$$

$$\leq h(Z_{1:p}) + \sum_{K=p+1}^{n} h(Z_K) Z_{K-p:K-1}) \quad (c)$$

$$\leq h(Z'_{1:p}) + \sum_{p+1}^{n} h(Z'_K | Z'_{K-p:K-1}).$$

$$\overset{(d)}{=} h(Z'_{1:p}) + \sum_{p+1}^{n} h(Z'_K | Z'_{1:K-1}) = h(Z'_{1:K}).$$

Inequalities.

(b) → chain rule inequality for entropy.

(C) → $h(A|B,C) \leq h(A|B)$

(d) → Markov property of Z'_K processes.

(a) → $D[f|\varphi] \triangleq \int f \log \frac{1}{\varphi} \geq 0$ always +ve, f is density of $\{X\}$

$\varphi = N(x|0,R) \to$ gaussian distribution

$R = E[xx^T] \to$ Covariance matrix

Conclusion: The p^{th} order gauss-markov process with Covariance $\alpha_{0:p}$ has higher entropy $h(Z'_{1:n})$ than any other process satisfying the autocorrelation Constraints, so that,

$$\lim_{n \to \infty} \frac{1}{n} \cdot h(X_{1:n}) \leq \lim_{n \to \infty} \cdot h(Z'_{1:n}) = h$$

for all stochastic process satisfying Constraints.

Equivalet Characterization:

$$X_n = \sum_{i=1}^{p} a_i X_{n-i} + U_n \qquad \{U_n\} \text{ iid}, \; U_n \sim N(0,\sigma^2)$$

$$\sigma_L = \mathbb{E}[X_n X_{n-L}]$$

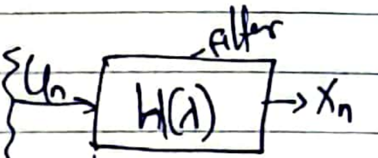
$$\sigma_2 = \overline{X X_{n-L}} = \sum_{j=1}^{p} a_j \cdot \underbrace{\overline{X_{n-i} X_{n-L}}}_{\sigma_{L-i}} + \overline{U_n \boxed{X_{n-1}}} \rightarrow U_{n-L}, U_{n-L+1} \cdots \qquad L > 0$$

$$\overline{U_n} \cdot \overline{X_{n-L}} = 0$$

$$\left.\begin{aligned} \sigma_L &= \sum_{j=1}^{p} a_j \sigma_{L-j} \quad L>0 \\ \sigma_{\infty} &= \sum_{ij} a_i a_j \sigma_{i-j} + \sigma_u^2 \end{aligned}\right\} \text{Yule-walker equations.}$$

$$\Gamma(\lambda) = \frac{\sigma_u^2}{\left|1 - \sum_{j=1}^{p} a_j e^{-i2\pi j\lambda}\right|^2}$$

Coz $H(\lambda) = \dfrac{1}{1 - \sum_{j=1}^{p} a_j e^{-i2\pi j\lambda}}$



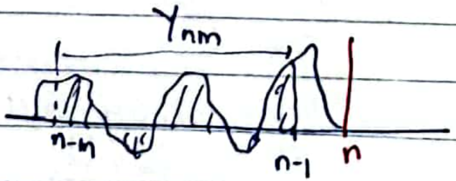
$$\underset{\text{spectrum o/p}}{\Gamma(\lambda)} = \underset{\text{spectrum input}}{\sigma_u^2} \cdot \underset{\text{response of filter.}}{|H(\lambda)|^2}$$

Fitting a Gauss Markov model to data (Linear prediction).
Suppose $\{X_n\}$ be any stationary process:

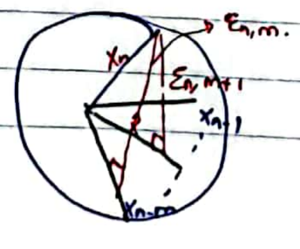
$$\varepsilon_n = X_n - L_{nm} \boxed{X_{n-1:n-m}} \rightarrow Y_{nm} \rightarrow \text{past of } X_n$$

L_{nm} = optimum projection of $X_n / X_{n-1:n-m}$

$L_{nm} = E[X_n \cdot Y_{nm}^T] R_{Y_{nm}}^{-1}$ from $L = E[X.Y^T]\,E[YY^T]$.



Var ~~of~~ (ε_n) decreases with m.



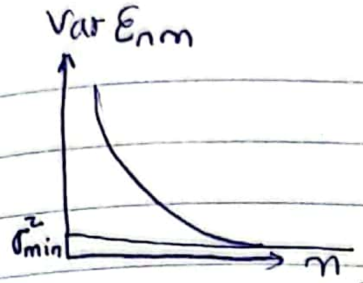
$$\varepsilon_{nm} = X_n - L_{nm} Y_n$$

$$\text{var } \varepsilon_{n,m-1} \leq \text{var } \varepsilon_{n,m}.$$

$$\varepsilon_{nm} \overset{\text{uncorrelated}}{\perp\!\!\!\perp} \{X_{n-1}, X_{n-2} \cdots\}$$

21/04 - 9-11
22/04 8-ves-1045

Gikhman . SKOROK - Random process
Ⓐ Todorovic → intro to stochastic process.



lec. Cont.

Max entropy - Estimation of PSD — optimisation

& after markov chain, Hidden markov (Rabinor + more)

Repetition

$\{X_1, \dots X_n\} \subseteq \{X_1, X_2 \dots\}$ stationary with $C \leq E[X_n X_{n \pm K}]$

$\downarrow$ Fourier T.

$\Gamma(\lambda) = \sum_{K \in \mathbb{R}} C_K e^{-2\pi i \lambda K}$

→ By using Max Entropy Criterion;

- Fit Gauss Markov model to available data → Autoregressive model

e.g. $X_n = \alpha_1 X_1 + \dots + \alpha_m X_{n-m} + \overset{\to \text{noise}}{u_n}$; when u_n is normal & iid → Gaussian noise

to fit the above data ; choose $\mathcal{E}$ which is function of α

$$\mathcal{E}_n(\alpha) = X_n - \sum_{K=1}^{n} \alpha_K X_{n-K} = X_n - \underline{\alpha}^T \cdot \underline{X}_{n-1:n-m}$$

minimize $\mathcal{E} \to E|\mathcal{E}_n(\alpha)|^2 \cong \frac{1}{N-m-1} \sum_{n=m+1}^{N} \{X_n^2 - 2\alpha^T X_n \cdot X_{n-1:n-m}$

$+ (\alpha^T X_{n-1:n-K})^2\}$ ↗ $\underline{\alpha}^T \underline{v}\, \underline{v}^T \alpha$

$$= \boxed{\frac{1}{N-m-1} \sum X^2} - \underbrace{2\alpha^T \frac{1}{N-m-1} \sum X_n \cdot X_{n-1:n-m}}_{\downarrow \text{covariance}} + \alpha^T \frac{1}{N-m-1} \sum \underbrace{X_{n-1:n-m} \cdot X^T_{n-1:n-m} \cdot \alpha}_{\downarrow \text{covariance of past}}$$

$\downarrow$ estimate

$\mathbb{E}[X^2]$

$\simeq \underbrace{\mathbb{E}[X_n \cdot X_{-n-1:n-m}]}_{\varphi}$

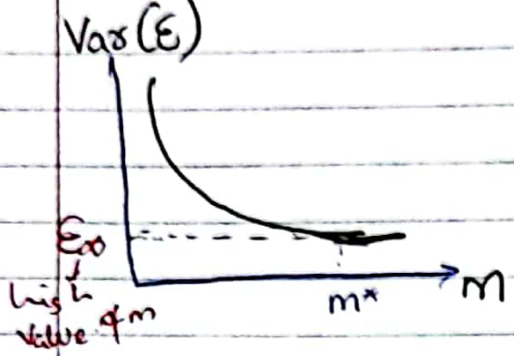
$\underbrace{\mathbb{E}[X_{n-1:n-m} \cdot X^T_{n-1:n-m}]}_{R}$

To minimize, best α

$$\hat{\underline{\alpha}}_m = R_m^{-1} \varphi_m$$

$$\text{Var}[\mathcal{E}_{n,m}] = \underbrace{\mathbb{E}[X_n^2] - \varphi_m^T R_m^{-1} \varphi_m}_{\hat{\sigma}_m^2}$$ → decreasing by m

french - 4 parts
- short resume of method
& matlab exercises.



→ $\varepsilon_{n,m} \perp\!\!\!\perp X_{n-1} \cdots X_{n-m}$

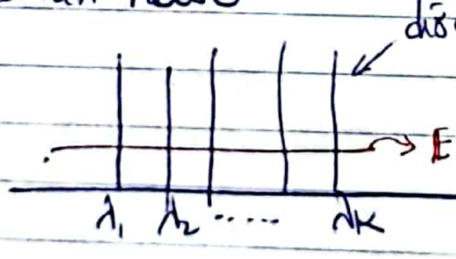
$\varepsilon_{n,m^*} \perp\!\!\!\perp \{X_{n-1} \cdots X_{n-2}\}$

$X_n = \alpha_1 X_{n-1} + \cdots \alpha_m X_{n-m} + U_n$

$$\hat{\Gamma}(\lambda) = \frac{\sigma^2}{1-}$$

Simulation.

line
3 N → No of observed samples
9 Signal → compared of sinusoids (sum of sinusoids) -
11 · Gaussian noise



X_n = Σ sinusoids at diff freqs + → sum of sinusoids

E[W_n^2) → var of white W_n → noise

λ → frequency

noise spectrum is constant.

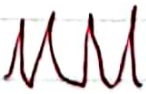
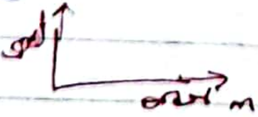
- $X_1, \ldots X_n$ → we want to get some-thing close to above.
- fit data X_n on gauss markov process of diffrent order.
- low order (m) - bad spectrum 
- high m → you get a better spectrum 

fig 3 

m<20, → maximum is low (blur).
m>60 → more clean spectrum.

$X_n \in \mathcal{S} \Rightarrow$ finite

~~Kalab~~ Algos

viterbi, BLTR.. Em Bau- welch.

$X_n \in$ Gaussian iid ~~[illegible]~~ $N(0,\sigma^2)$

$\downarrow$

Kalman filter

Markov Process & chains.

- for optimatization
- to get max, min of complex process

$Y_n = f(X_n, W_n)$ ↗ X_n is hidden in noise

$\hat{X}_n = X_n$

Asymptotic Behaviour of Markov chains (Ergodicity).

→ long run

channel

$$U_n \in \{-1, \pm 1\} \rightarrow \boxed{h_0, h_1} \rightarrow Z_n = h_0 U_n + h_1 U_{n-1} \xrightarrow{\downarrow W_n} Y_n = Z_n + W_n \pm \tfrac{1}{2}$$

Prob.

↙ Independ.

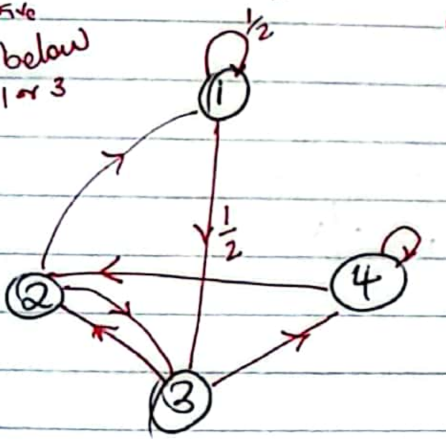
$$X_n = \begin{bmatrix} U_n \\ | \\ U_{n-1} \end{bmatrix} \in \left\{ \begin{bmatrix} + \\ + \end{bmatrix}, \begin{bmatrix} \textcircled{+} \\ - \end{bmatrix}, \begin{bmatrix} - \\ + \end{bmatrix}, \begin{bmatrix} - \\ - \end{bmatrix} \right\} = \mathcal{S}$$

1 2 3 4 → 4 possible states.

→ last two sent samples were [+1, +1], next two possible states

next

have next + below hence 1 or 3



(*)

$$P(X_n = j \mid X_{n-1} = i) \quad i,j \in \mathcal{S} = A_{ij} \rightarrow \text{transition matrix.}$$

→ conditioning its last value ×tories markov chain

$$P[X_n = i_n \mid X_0 = i_0 \ldots X_{n-1} = i_{n-1}] = P[X_n = i_n \mid X_{n-1} = i_{n-1}]. \rightarrow \text{matrix}$$

present ↓ future →

For the above

$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

$$P[X_0] \rightarrow P_i = P(X_0 = i] \rightarrow \text{vector}$$

$\mathbb{P}[X_n = j] = \sum_{i} P[X_n = j, X_{n-1} = i]$, $\quad j \in \$$

$= \sum_{i \in \$} P[X_n = j \mid X_{n-1} = i]\, P[X_{n-1} = i]$.

$P_n(j) = \sum_{i \in \$} P_{n-1}(i) \cdot A_{ij}$ → vector → stochastic vector

→ stochastic matrix.

Current $P_n = P_{n-1} \cdot A$ $\quad 0 \le A_{ij} \le 1$

$= P_0 \cdot A^n$. $\quad \sum_{i} A_{ij} = 1$ → sum of all future

P_0 → initial prop

Ergodic case: $\forall P_0$, $P_0 A^n \xrightarrow{\text{Converge}} P_\infty$ e.g. distribution of link becomes same; forget the initial prob

→ first slides

→ $\forall \pi_0 : \pi_0 A^n \rightarrow \dot{\pi}_\infty$ → Discuss this for $n \to \infty$

NB $\underset{\sim}{P_n(i)} = P[X_n = i]$. → $i = 1 : T$

$\pi_n = \pi_{n-1} \cdot A = \pi_0 A^n$

hence $\pi_\infty A = \pi_\infty$

i.e. π_∞ is an invariant distribution

Def: Matrix A is ergodic if there exist ($\exists$) s such that $A_{ij}^{(s)}$

i.e. matrix $A^{(s)}$ at element i & j > 0 [illegible]

- π_i is ergodic of inv

Ergodic Theorem: If A is ergodic, an invariant unre

$A_{ij}^{[n]} = \pi_j$
$n \infty$

Example: Metropolis algorithm