08/30/19

AMS 131 Lecture 15

The set of values a Markov chain can take on is called its <u>state space</u> S, finite or infinite
(Can also have Markov chains unfolding in continuous time, ex: X_t = stock price at time t
t = time)

1. $(X_1, X_2 \ldots)$ finite Markov chain $\rightarrow$

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdot P(X_2 = x_2 \mid X_1 = x_1) \cdot P(X_3 = x_3 \mid X_2 = x_2) \ldots \cdot P(X_n = x_n \mid X_{n-1} = x_{n-1})$$

suppose you have a finite Markov chain with k possible states numbered $1, \ldots, k$
(k integer ≥ 2) $\rightarrow \{P(X_{n+1} = j \mid X_n = i\}$ $i = 1, \ldots, k$ $n = 1, 2 \ldots$
is called the transition distribution of the M.C.

If $P(X_{n+1} = j \mid X_n = i)$ is the same for all n, the transition distribution is said to be stationary (time-homogeneous)

↳ if so, $p_{ij} \triangleq P(X_{n+1} = j \mid X_n = i)$ completely characterize the Markov chain's behavior

$$P = \underset{\text{state}}{\text{from}} \quad \overset{\text{to state}}{\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ \vdots & & \ddots & \cdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}}$$

transition matrix P

row must add up to $1 \rightarrow$ chain goes <u>somewhere</u>
all elements $0 < p_{ij} < 1$ (probabilities)

A square matrix $\underset{k \times k}{P}$ with non-negative entries and all row sums $= 1$ is a <u>stochastic matrix</u>.

Gene inheritance is markovian, the genetic makeup for you is only determined by your parents., ancestors are irrelevant because they essential make up your parents.

Say a gene of interest has alleles A and a, then a state in the Markov chain is in the form:

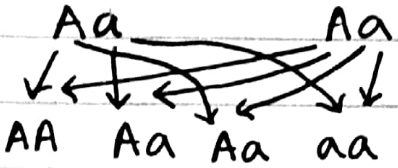
{allele 1 from parent 1 allele 2 from parent 1, allele 1 from parent 2 allele 2 from parent 2}

ex: {Aa, Aa}

ignoring order (not important in inheritance): 6 states

{AA, AA}, {AA, Aa}, {AA, aa}, {Aa, Aa}, {Aa, aa}, {aa, aa}

one possible inheritance sequence



Offspring gets A or a from each parent (independent), each with probability $\frac{1}{2}$

transition matrix

from \ to	{AA,AA}	{AA,Aa}	{AA,aa}	{Aa,Aa}	{Aa,aa}	{aa,aa}
{AA,AA}	1	0	0	0	0	0
{AA,Aa}	$\frac{1}{4}$	$\frac{1}{2}$	0	$\frac{1}{4}$	0	0
{AA,aa}	0	0	0	1	0	0
{Aa,Aa}	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{16}$
{Aa,aa}	0	0	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
{aa,aa}	0	0	0	0	0	1

Random walk: you're watching a particle move on the integers $S = \{\ldots -2, -1, 0, 1, 2 \ldots\}$ over time

- Wherever it is at time n it moves
  - left 1 unit with prob p_1
  - stays with prob p_2
  - right 1 unit with prob p_3

$0 < p_i < 1$

$\sum_{i=1}^{3} p_i = 1$

this is a markov chain because its future position only depends on where it is, not its previous.

AMS 131

Leture 15 (cont.)

from \ to	...	-2	-1	0	1	2	...
...	...	...	...	...	...	...	...
-2	...	p_2	p_3	0	0	0	...
-1	...	p_1	p_2	p_3	0	0	...
0	...	0	p_1	p_2	p_3	0	...
1	...	0	0	p_1	p_2	p_3	...
2	...	0	0	0	p_1	p_2	...
...	...	...	...	...	...	...	...

transition matrix $= \underline{\underline{P}}$

band matrix: only nonzero entries are on the main diagonal and one diagonal away from the main

- 3 diagonals: tridiagonal

all the main diagonal entries are the same p_2

one above: p_3, one below: p_1

Start the random walk, at 0 let it go, where is it likely to be at time n, n large

Suppose: $(p_1, p_2, p_3) = (0.1, 0.3, 0.6)$

will eventually drift to $+\infty$

$(0.5, 0.25, 0.25)$ drift to $-\infty$

$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) = ?$ as $n \to \infty$, each integer is visited infinitely many times and the expected time for it to return to 0 is also infinite $\rightarrow$ "too much freedom"

We can restrict the bounds to: for ≥ 1

$S = \{-K, -K+1, \ldots -1, 0, 1, \ldots, K-1, K\}$

at the boundary

$K-1$ K $K+1$ not possible

so we can wrap around: if you move to $(K+1)$, go to $-K$ and if move to $-(K+1)$, go to K

transition matrix, $k=2$

$$\begin{array}{c|ccccc} & -2 & -1 & 0 & 1 & 2 \\ \hline -2 & p_2 & p_3 & 0 & 0 & p_1 \\ -1 & p_1 & p_2 & p_3 & 0 & 0 \\ 0 & 0 & p_1 & p_2 & p_3 & 0 \\ 1 & 0 & 0 & p_1 & p_2 & p_3 \\ 2 & p_3 & 0 & 0 & p_1 & p_2 \end{array}$$

from (rows) / to (columns)

another idea is to make the boundaries impossible to cross, hard stop.

$P_{ij}^{(m)}$ = P(chain moves from i to j in m steps)
$= P(X_{n+m} = j \mid X_n = i)$

finite Markov chain with stationary transition distributions and transition matrix $\underline{P} \rightarrow P_{ij}^{(m)}$ is the (i,j) entry of the matrix P^m, which is called the m-step transition matrix.

{AA, AA} and {aa, aa} has the property that once the chain is it that state, it cannot go anywhere else - state i when $p_{ii} = 1$
any state with $p_{ii} = 1$, is called an absorbing state

In the genetic chain, 1,2,4 have positive probability of moving to state 1 in 2 steps, and the same for moving to state 6 in 2 steps.

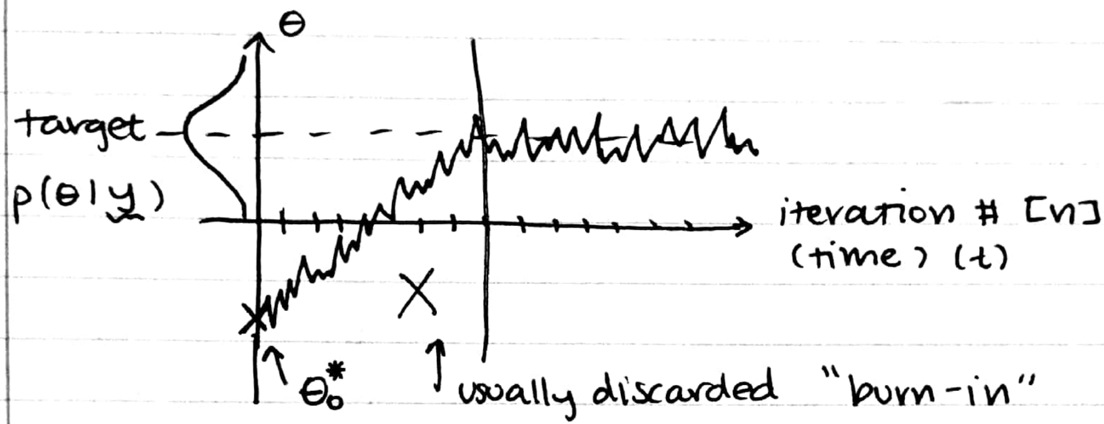
Markov chains that settle down to a single, stable, long run distribution (equilibrium distribution) are called steady state (stationary dist). ← confusing with time-homogeneous

Where should the Markov chain start?
- either initialize it at a deterministic value or make a random draw from the initial distribution

AMS 131

Lecture 15 (cont.)

Any vector $\underset{\sim}{v}$ of non negative numbers that add up to 1 is called a probability vector - any such vector whose components specify that a Markov chain will be in each possible state at time 1 is the initial distribution of the chain.



After 1 time step (iteration) the probability dist. over the Markov chain's possible states is $\underset{\sim}{v}$, after 2 iterations, the chain's dist is: $\underset{\sim}{v}_{1 \times k} \underset{=}{P}_{k \times k}$, after $(m+1)$, $\underset{\sim}{v} \underset{=}{P}^m$, looking for it to converge to a unique dist. as $m \to \infty$, resulting in its equilibrium distribution

If we choose $\underset{\sim}{v} \underset{=}{P} = \underset{\sim}{v}$, then
$$\underset{\sim}{v} \underset{=}{P}^2 = (\underset{\sim}{v} \underset{=}{P}) P = \underset{\sim}{v} \underset{=}{P} = \underset{\sim}{v}$$
$$\underset{\sim}{v} \underset{=}{P}^3 = (\underset{\sim}{v} \underset{=}{P}^2) \underset{=}{P} = \underset{\sim}{v} \underset{=}{P} = \underset{\sim}{v}$$
$$\lim_{m \to \infty} \underset{\sim}{v} \underset{=}{P}^m = \underset{\sim}{v}$$

Markov chain with transition matrix $\underset{=}{P}$ → only probability vector $\underset{\sim}{v}$ such that $\underset{\sim}{v} \underset{=}{P} = \underset{\sim}{v}$ is an equilibrium dist. for the chain

- under additional conditions on $\underset{=}{P}$, such an eq. dist. will be unique

Lecture 15 (cont.) AMS 131

$\underset{1\times K}{\underset{\sim}{V}}\ \underset{K\times K}{P} = \underset{1\times K}{\underset{\sim}{V}}$ eigen-analysis of P

$\underset{K\times K}{A}\ \underset{K\times 1}{\underset{\sim}{X}} = \lambda \underset{\sim}{X}$ ← an eigenvector
↖ an eigenvalue
} both "right"

"left" eigenvector $VP = \overset{1}{V}$ ← eigenvalue of 1

Given a square matrix $\underset{K\times K}{P}$, any vector $\underset{K\times R}{\underset{\sim}{V}}$ satisfying $\underbrace{\underset{K\times K}{P}\ \underset{\sim}{V}_{R_1}}_{K\ 1} = \lambda_R \underbrace{{}_K\underset{\sim}{V}_{R_1}}_{K\ 1}$ is called a right eigenvector of P with eigen value λ_R and any vector ${}_1\underset{\sim}{V}_{L_K}$ satisfying ${}_1\underset{\sim}{V}_{L}\ \underset{K\times K}{P} = \lambda_L\ {}_1\underset{\sim}{V}_{L_K}$ is called a left eigenvector with eigenvalue λ_L

So given a transition matrix $\underset{K\times K}{P}$ for a Markov chain, an equilibrium dist. can be found by computing the left eigenvector ${}_1\underset{\sim}{V}_{L_K}$ where the eigenvalue is 1. notice if:

${}_1\underset{\sim}{V}_{L_K}\ \underset{K\times K}{P} = {}_1\underset{\sim}{V}_{L_K}$ then $({}_1\underset{\sim}{V}_{L_K}\ \underset{K\times K}{P})^T = {}_1\underset{\sim}{V}_{L_K}^T$

$= \underset{K\times K}{P^T}\ \underset{\sim}{V}_L{}^T_1 = {}_K\underset{\sim}{V}_L{}^T_1$

so we can eigendecompare $\underline{P}^T$ instead of $\underline{P}$

$\begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}$ and $\begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix}$ correspond to the absorbing states (of genetics problem), this suggests there's a family of eq. dist. of the form:

$(p, 0, 0, 0, 0, 1-p)^T$ for $0 \le p \le 1$

If you can find an integer $m \ge 1$ such that every element of $\underline{P}^m$ is strictly positive, then $\lim_{n\to\infty} \underline{P}^n$ is a matrix with all rows equal to the unique stationary dist. $\underset{\sim}{V}$ and no matter what the initial dist. is, its dist after n steps converges to $\underset{\sim}{V}$ as $n \to \infty$

AMS 131

Lecture 15 (cont.)

case study: stock options

companies can give signing bonuses often in the form of stock options: an opportunity to buy N shares of the company one year from now at a (known) price., S

If you think the stock will rise = sell at profit.

X: price of stock 1 yr from now. ← not known

↳ for now, discrete with two values: $x_1 < S$ and $x_2 > S$ let:

$p = P(X = x_2)$ the prob that stock value will rise

Y = value of option for one share at \$ S

If ($X = x_1 < S$), option is worthless and $Y = 0$

If ($X = x_2 > S$) then the option is worth ($x_2 - S$) thus $Y = h(X) = \begin{cases} 0 & \text{if } X = x_1 \\ (x_2 - S) & x_2 \end{cases}$

- Can compare stock option to investing in a bond that pays α%/year →
- A fair measure of the worth of the option would be the present value of Y, defined to be the number c such that $E(Y) = (1+\alpha)\cdot c$

$E(Y) = 0\cdot(1-p) + (x_2 - S)p = (x_2 - S)p$

$(1+\alpha)\cdot c = (x_2 - S)p$ comparing stock options to bonds

need p → $c = \left(\dfrac{x_2 - S}{1+\alpha}\right)p$

assume the present value of X is equal in expectation to the current value of the stock.

$\left(\begin{array}{l}\text{buying 1 share}\\ \text{and hold 1 yr.}\end{array}\right) = \left(\begin{array}{l}\text{investing same}\\ \text{amount into bond}\end{array}\right)$

$E(X) = (1+\alpha)\cdot S$

but $E(X) \doteq p\cdot x_2 + (1-p)x_1 \stackrel{\wedge}{=} (1+\alpha)\cdot S$

$p = \dfrac{x_1 - (1+\alpha)\cdot S}{x_1 - x_2} = \dfrac{(1+\alpha)S - x_1}{x_2 - x_1}$

So, fair price c of an option to buy one share : $c = \left(\frac{x_2 - S}{1+\alpha}\right)\left(\frac{(1+\alpha) \cdot S - x_1}{x_2 - x_1}\right)$

$S = \$200$ $\alpha = 0.04$ ← realistic

$x_1 = \$180$ downside \$20 (−10%) $c \doteq \$20.19$

$x_2 = \$260$ upside \$60 (+30%) ~10% of current

c = risk neutral price of the option

Under the assumptions we made, you could now sell the option today (if you had it) at a fair price ≐ \$20 (options trading)

An investment that allows people to buy or sell an option on a security is called a derivative

case study: building a good portfolio.

You have \$I to invest to build a portfolio, optimal collection for companies A, C + bonds B

This depends on how much A, C will change

$R_A \triangleq$ rate of return (change in price) for comp. A

$R_C \triangleq$ " C "

your portfolio: (S_A shares of A stock, S_C shares of C stock, \$$S_B$ worth of bonds)

1 year return: $\underbrace{S_A R_A + S_C R_C}_{\text{random}} + \$r S_B$

fixed random fixed

$E(R_A) = \mu_A$ $V(R_A) = \sigma_A^2$ (S_A, S_C, S_B) : total current value = \$I

$E(R_C) = \mu_C$ $V(R_C) = \sigma_C^2$

P_A = current price of A, P_C = " of C

constraint : $S_A P_A + S_C P_C + S_B = I$

Pretend R_A, R_C independent : $E(S_A R_A + S_C R_C + r S_B) = S_A \mu_A + S_C \mu_C + r S_B$

$V(S_A R_A + S_C R_C + r S_B) = S_A^2 \sigma_A^2 + S_C^2 \sigma_C^2$

↳ we want: expected high, variance low (rare)

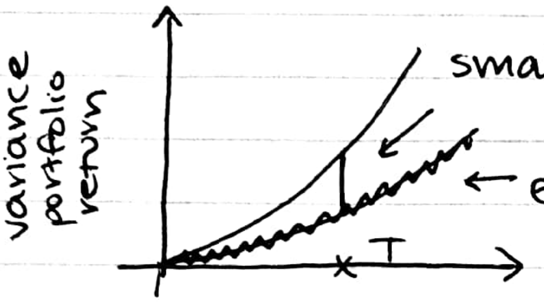
In practice need to sacrifice some expected return to keep downside risk small.

AMS 131 | Lecture 15 (cont.)

How should you optimally make this trade off?

optimize over (U, V)

1. optimize $pU + (1-p)V$ ✓ or vice versa
2. optimize on U subject to constraint on V



Suppose you set target \$T for your expected return $E(\text{return}) = T$ with smallest variance = best

↳ minimal variance portfolios are efficient

for fixed: P_A, P_C, T, I, r , $\underbrace{\mu_A, \mu_C, \sigma_A^2, \sigma_C^2}_{\text{need to speculate}}$

find (S_A, S_C, S_B) to

minimize $V = S_A^2\sigma_A^2 + S_C^2\sigma_C^2$ subject to

$E = S_A\mu_A + S_C\mu_C + rS_B = T$ and

$S_AP_A + S_CP_C + S_B = I$

instead of first derivatives, use Lagrange multipliers

$I = \$100{,}000$ $T = \$7{,}000$ (7% reasonable return)

$V = 2.55 \times 10^7\ \$^2$, $SD \doteq \$5050$

If dist. of return follows bell curve, only 8% chance of losing money.