① HW4

2-3-15

Bivariate Normal Distributions and Gibbs Sampling

Introduction: If we let $X \sim N(\mu, \sigma^2)$ then we are dealing with a univariate distribution that has two parameters

- If X and Y are both standard normal r.v.s with correlation $\rho = 0$ then their joint density is given by:

$$f(x,y) = \frac{1}{\sqrt{2\pi}} e^{-(x^2+y^2)/2} \quad \text{for } x \in \mathbb{R}, y \in \mathbb{R}$$

This is a bivariate distribution with 5 parameters $\mu_x = 0$ $\mu_y = 0$ $\sigma_x^2 = 1$ $\sigma_y^2 = 1$ $\rho = 0$

- If X and Y are both standard normal with $\rho \in (-1, 1)$, then the joint density is given by:

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-(x^2+y^2-2\rho xy)/2(1-\rho^2)} \quad \text{for } x \in \mathbb{R}, y \in \mathbb{R}$$

Marginal density for Y

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

$$e^{-\frac{(x^2 - 2\rho xy + \rho^2 y^2 + (1-\rho^2)y^2)}{2(1-\rho^2)}}$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{\frac{-y^2}{2(1-\rho^2)}} e^{-(x^2-2\rho xy)/2(1-\rho^2)}\, dx$$

complete the square

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot e^{-y^2/2(1-\rho^2)} \int_{-\infty}^{\infty} e^{-(x-\rho y)^2/2(1-\rho^2)} \cdot e^{\frac{\rho^2 y^2}{2(1-\rho^2)}}$$

because subtracted $e^{\rho y}$ to complete the square

Marginal integrate all variables not interested in out

②

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot e^{-y^2/(2(1-\rho^2))} \cdot e^{\rho^2 y^2/2(1-\rho^2)} \int_{-\infty}^{\infty} e^{-(x-\rho y)^2/2(1-\rho^2)}$$

$$= \underbrace{\frac{1}{2\pi\sqrt{1-\rho^2}} e^{-y^2/2} \cdot \sqrt{2\pi(1-\rho^2)}}_{\text{from cancellation}} \boxed{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-(x-\rho y)^2/2(1-\rho^2)} dx} = 1$$

to make it in the form of a normal dist. $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

$\mu = \rho y$

from cancellation

$$\frac{1}{\sqrt{2\pi}} e^{-y^2/2} \quad y \in \mathbb{R}$$

Marginal density for X is same but in terms of X

Full conditional Distribution of Y $\left[-(y^2 - 2\rho xy + \rho^2 x^2) - x^2(1-\rho^2)\right]$

proportional

$$f(y|x,\rho) \propto \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-(y-\rho x)^2/2(1-\rho^2)} e^{-x^2/2} \quad \text{for } y \in \mathbb{R}$$

remember though - it is proportional not equal

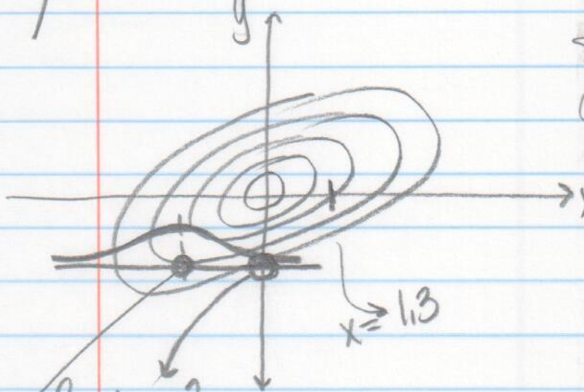
write down joint def. considering all var fixed except

$$y|x,\rho \sim N(\underbrace{\rho x}_{E(Y|X)}, \underbrace{(1-\rho^2)}_{var(Y|X)})$$

$x|y,\rho \sim N(\rho y, (1-\rho^2)) \leftarrow$ Full conditional Dist. of X

$f(x,y)$

$\rho > 0$ maybe 0.6



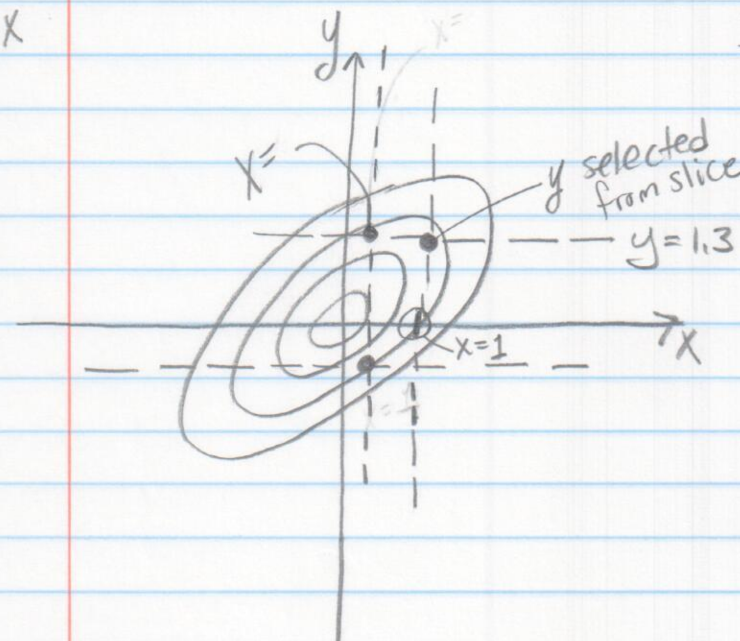
conditional is centered at x

<u>marginal</u>

squash into y-axis or x-axis and get 2-D bell shape ~ N centered <u>at 0</u>

<u>full</u> <u>conditional</u>

slice at 1.3



Initialize $x_1 = 1$

- plug into full conditional for y and sample a value of y → sample 1.3

new x from y

- plug into full conditional for x and sample a value of x → sample

new x ...

- plug into full cond. for y

$$\begin{array}{r} \text{Iteration } 1 \\ 2 \\ \vdots \\ 10000 \end{array} \begin{bmatrix} x_1 , & y_1 \\ x_2 , & y_2 \\ \vdots & \vdots \\ x_{10000} , & y_{10000} \end{bmatrix}$$

x – y pairs are a sample from $f(x,y)$

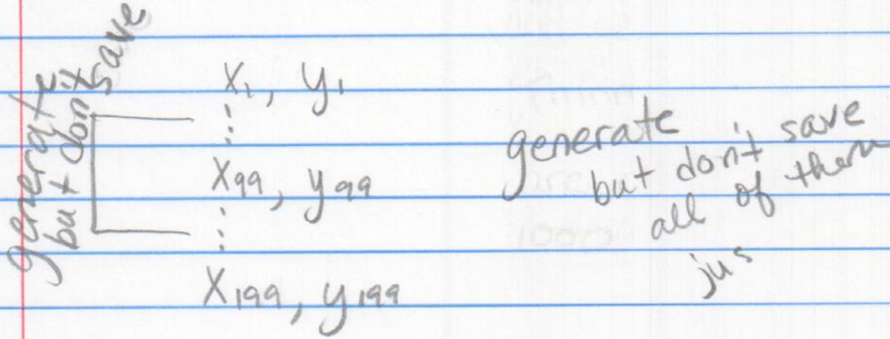
all xs sample from marginal for x

all ys are a sample from marginal for y

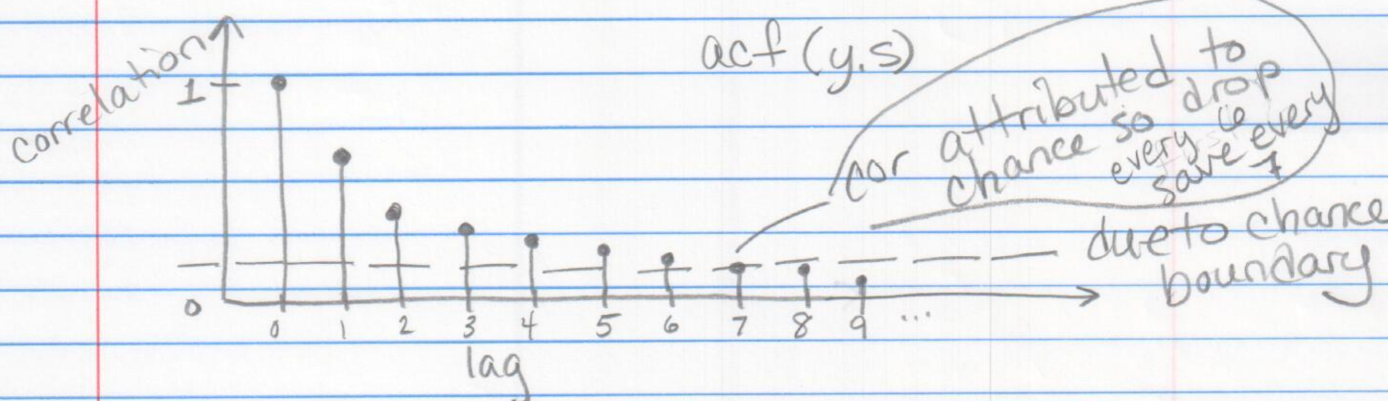
Auto Correlation and Thinning

thinning

Above realizations are not independent
skip every 100, 300, 1000 - how do we know how often to thin



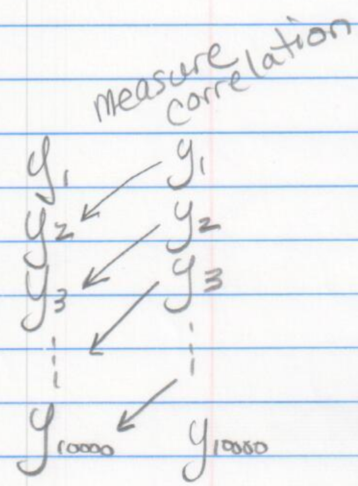
An autocorrelation plot is a diagnostic to help determine the lag necessary to achieve independent realizations



Lag 0: cor(y.s, y.s)

Lag 1: cor(y.s[1:(n-1)], y.s[2:n])

Lag 2: cor(y.s[1:(n-2)], y.s[3:n])

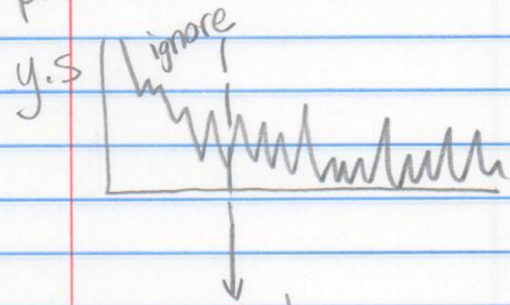


Burn-in

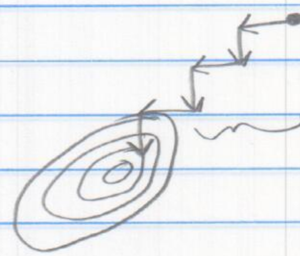
$X = 8$

ts.plot(y.s)

Trace plot



start saving here



chop off to adjust for obsurd values of X

CPMA 573 — Homework #4

Exercise 1: Recognizing burn-in and autocorrelation. Let X and Y have jointly normal density $f(x,y)$ given by:

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp\left\{-(x^2+y^2-2\rho xy)/(2(1-\rho^2))\right\} .$$

Write a program that uses Gibbs sampling to produce realizations from this joint density, where $\rho = 0.75$. Your program should be flexible enough to accommodate a choice of

- Initial values x_0 and y_0.
- Number of desired (x_i, y_i) realizations.
- Number of iterations to skip (lag) in order to avoid serial correlation.

Use this program to complete the following:

a. Saving every iteration, simulate 500 realizations from starting values $x_0 = 80$ and $y_0 = 80$. Plot the first 50 realizations in Splus, and connect-the-points in the order you generated them using the "arrows" command:

```
plot(x[1:50],y[1:50])
arrows(x[1:49],y[1:49],x[2:50],y[2:50],size=0.1,open=T)
```

b. Construct the same plot as in **a**, only this time, use all 500 realizations. This type of plot gives an idea of the necessary "burn-in" iterations, for given starting values.

c. When using $x_0 = 80$ and $y_0 = 80$ as initial values, how many iterations (roughly) does it take for the Gibbs sampler to converge to its stationary distribution? Reference the two trace plots (Use `ts.plot` once for the x realizations, and once for the y. Try using the argument `xlab=''Iterations''` to change the x-axis label to something more appropriate.)

d. Now create two autocorrelation plots (`acf`): One for the 500 x-values, and one for the 500 y-values. Based on these plots, what lag should be used to eliminate autocorrelation?

e. Using the lag from your answer to **d**, use your Gibbs sampling program to draw 10000 independent realizations from the marginal density of X. Confirm, using the mean, variance, and quartiles, that these realizations are indeed from a standard normal density.

Exercise 2: The influence of ρ. Use your Gibbs sampling program from Exercise 1 to generate 15000 independent (x, y) realizations. Then, use these realizations to find the following:

a. $P(X > 0 \text{ and } Y > 0)$, for $\rho = \{-0.9, -0.8, \ldots, 0.8, 0.9\}$. Then plot these probabilities as a function of ρ. Set the vertical axis limits to $(0, 1)$.

b. $P(X < 0 \text{ and } Y > 0)$, for $\rho = \{-0.9, -0.8, \ldots, 0.8, 0.9\}$. Then plot these probabilities as a function of ρ. Set the vertical axis limits to $(0, 1)$.

c. $P(X > 0)$, for $\rho = \{-0.9, -0.8, \ldots, 0.8, 0.9\}$. Then plot these probabilities as a function of ρ. Set the vertical axis limits to $(0, 1)$.