LECTURE NOTES

CHANDLER

§ 1 | An Application

We operate a cufflink machine:

The weights of the two ends, M_1 and M_2 are *i.i.d* $N(5, 10)$. We want to only keep cufflinks that are balanced. That means $|M_1 - M_2|$ is small.

A. It's easy to weigh both ends, $M_1 + M_2$. Does this help identify the bad cufflinks?

No. $M_1 + M_2$ and $M_1 - M_2$ are independent. This is because:

$$\mathrm{cov}[M_1 + M_2, M_1 - M_2] = \mathrm{var}[M_1] - \mathrm{var}[M_2] = 0\,.$$

B. What is the distribution of M_1 among cufflinks with $M_1 + M_2 \approx 12$?

Let $A = M_1 + M_2$ and $B = M_1 - M_2$, then $M_1 = \frac{A+B}{2}$. From Part A we know A and B are independent.

$$(M_1 | A = 12) \stackrel{d}{=} \frac{12 + B}{2} \sim N(6, 2)$$

This result follows from the fact that:

$$B \sim N(0, 2)$$

since $B = M_1 - M_2$.

§ 2 | Conditional Distributions

2.1 Lemma (Conditional Distributions) Let $\underset{n\times n}{X}$ and $\underset{m\times m}{Y}$ be jointly Gaussian with mean zero and covariance:

$$\begin{bmatrix} \underset{n\times n}{\Sigma_{XX}} & \underset{m\times n}{\Sigma_{XY}} \\ \underset{n\times m}{\Sigma_{XY}^{\mathrm{T}}} & \underset{m\times m}{\Sigma_{YY}} \end{bmatrix}$$

NOTE:

Then,

$$\mathbb{E}[X|Y] = \Sigma_{XY}\Sigma_{YY}^{-1}Y$$

and

$$(X|Y = y) \sim N(\Sigma_{XY}\Sigma_{YY}^{-1}, \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}})$$

Date: 03 October 2018.

Note that if Σ_{YY} is not invertible, we may use the Moore-Penrose inverse. That is, the above equations then remain true if the inverse used is the Moore-Penrose generalized inverse. Further, note that $\Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}}$ is the Schur complement of our covariance matrix.

Is this line correct?

Proof. Let $A = \Sigma_{XY}\Sigma_{YY}^{-1}Y$ and $X = A + B$. Claim: B is independent of Y.

$$\operatorname{cov}[X - \Sigma_{XX}\Sigma_{YY}^{-1}Y, Y] = \operatorname{cov}[X, Y] = \operatorname{cov}[X, Y] - \Sigma_{XY}\Sigma_{YY}^{-1}\operatorname{cov}[Y, Y] = 0\,.$$

Now we can say:

$$(X|Y = y) \overset{d}{=} \Sigma_{XY}\Sigma_{YY}^{-1}y + B$$

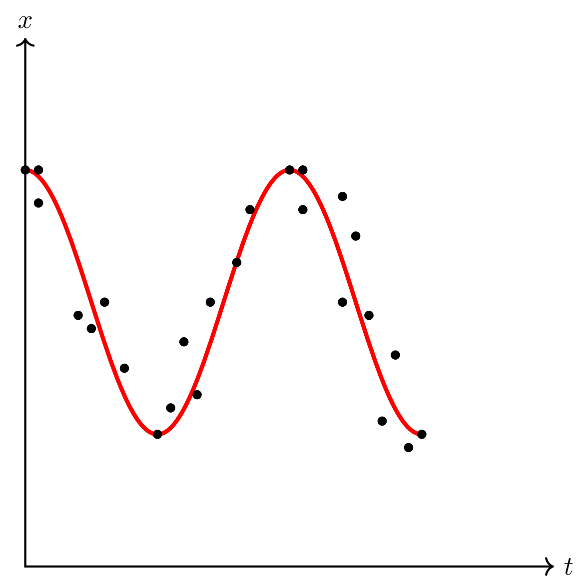
To confirm that this is the same distribution as before we need to calculate:

$$\begin{aligned}\operatorname{cov}[B, B] &= \operatorname{cov}[X - \Sigma_{XY}\Sigma_{YY}^{-1}Y, X - \Sigma_{XY}\Sigma_{YY}^{-1}Y] \\ &= \operatorname{cov}[X, X] - \operatorname{cov}[X, Y]\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}} - \Sigma_{XY}\Sigma_{YY}^{-1}\operatorname{cov}[Y, X] + \Sigma_{XY}\Sigma_{YY}^{-1}\operatorname{cov}[Y, Y]\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}} \\ &= \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}}\end{aligned}$$

Thus, $(X|Y = y) \sim N(\Sigma_{XY}\Sigma_{YY}^{-1}, \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^{\mathrm{T}})$. □

§ 3 | Gaussian Process Regression

Suppose we have a random curve that is drawn from a Gaussian process. We observe the value of this process at a few noisy points.



Using this information we can find out more about this curve, including confidence intervals about points.

But, first we'll look more at the Brownian Bridge. Recall (for $0 \le t < 1$),

$$U_t = (B_t | B_1 = 0) \overset{d}{=} B_t - tB_1$$

Check this,

$$\operatorname{cov}[B_1, B_t - tB_1] = t - t \cdot 1 = 0$$

Recall,

$$U_t = B_t - Z_{0,0}\int_0^t h_0(s)ds = Z(\mathbf{1}_{[0,t)})) - tZ_{0,0} = \sum_{n\geq 1}\sum_{k=1}^{2^n} Z_{n,k}\int_0^t h_{n,k}(s)ds$$

What is the distribution of $U_t - U_s$? (for $s < t$)

$$U_t - U_s = \sum_{n\geq 1}\sum_{k=1}^{2^n} Z_{n,k}\int_s^t h_{n,k}(s)ds = Z(\mathbf{1}_{[s,t)}) - (t-s)Z_{0,0}$$

So,

$$\operatorname{var}[U_t - U_s] = \operatorname{var}[Z(\mathbf{1}_{[s,t)})] - 2(t-s)\cdot\operatorname{cov}[Z(\mathbf{1}_{[s,t)}), Z(\mathbf{1}_{[0,1)})] + (t-s)^2\cdot\operatorname{var}[Z(\mathbf{1}_{[0,1)})]$$

Note that,

$$\operatorname{var}[Z(\mathbf{1}_{[s,t)})] = \int_0^1 \mathbf{1}_{[s,t)}(u)\cdot\mathbf{1}_{[s,t)}(u)du = t - s$$

So, we'll have

$$\operatorname{var}[U_t - U_s] = (t-s) - 2(t-s)^2 + (t-s)^2 = (t-s)[1-(t-s)]$$

Thus,

$$U_t - U_s \sim N\Big(0, (t-s)[1-(t-s)]\Big)$$