CMSC 25025
Homework 1

Reid McIlroy-Young

April 6, 2017

1 Statistical refresher

1.a

$$
\begin{aligned}
F(X) &= \int_{-\infty}^{X} f_x(t)dt \\
\Rightarrow f_x &= \frac{dF}{dX}
\end{aligned}
$$

$$
\begin{aligned}
X &= F^{-1}(U) \\
F(X) &= F(F^{-1}(U)) \\
F(X) &= U \\
\Rightarrow X &\sim F
\end{aligned}
$$

1.b

$$
X, Y \sim Uniform(0,1) \quad (1)
$$

$$
f_{X,Y}(x,y) = 1 \quad (2)
$$

$$
\begin{aligned}
Z &= X - Y \\
F_Z(z) &= P(Z \leq z) \\
&= P(X - Y \leq z) \\
&= \begin{cases} \int_0^{1+z} \int_{y-z}^{1} dxdy & 1 \leq z \leq 0 \\ 1 - \int_z^1 \int_0^{y-z} dxdy & 0 < z \leq -1 \\ 0 & else \end{cases} \\
&= \begin{cases} \int_0^{1+z} 1 - y - zdy & 1 \leq z \leq 0 \\ 1 - \int_z^1 y - zdxdy & 0 < z \leq -1 \\ 0 & else \end{cases} \\
&= \begin{cases} 1 + z - \frac{(1+z)^2}{2} - z(1+z) & 1 \leq z \leq 0 \\ 1 - (\frac{(1)^2}{2} - z - \frac{(z)^2}{2} + z^2) & 0 < z \leq -1 \\ 0 & else \end{cases} \\
&= \begin{cases} \frac{(z)^2}{2} + z + \frac{1}{2} & 1 \leq z \leq 0 \\ -\frac{(z)^2}{2} + z + \frac{1}{2} & 0 < z \leq -1 \\ 0 & else \end{cases} \\
\Rightarrow f_z(z) &= \begin{cases} z + 1 & 1 \leq z \leq 0 \\ -z + 1 & 0 < z \leq -1 \\ 0 & else \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
Z &= min\{X, Y\} \\
F_Z(z) &= P(Z \leq z) \\
&= P(min\{X, Y\} \leq z) \\
&= 1 - P(min\{X, Y\} \geq z) \\
&= 1 - P(X \geq z)P(Y \geq z) \\
&= 1 - (1 - z)(1 - z) \\
&= 2z - z^2
\end{aligned}
$$

1.c

$$
\begin{aligned}
x &\sim N(0,1) && (3)\\
Y &= e^X && (4)\\
u^{-1}(Y) &= log(Y) && (5)\\
\frac{du^-1}{dY} &= \frac{1}{Y} && (6)\\
\Rightarrow f_Y(y) &= \frac{1}{y\sqrt{2\pi}} e^{\frac{log^2(y)}{2}} && (7)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}(Y) &= \int_{-\infty}^{\infty} y \frac{1}{y\sqrt{2\pi}} e^{\frac{log^2(y)}{2}} dy\\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{log^2(y)}{2}} dy\\
&= e^{\frac{1}{2}}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}(Y^2) &= \int_{-\infty}^{\infty} y^2 \frac{1}{y\sqrt{2\pi}} e^{\frac{log^2(y)}{2}} dy\\
&= \int_{-\infty}^{\infty} \frac{y}{\sqrt{2\pi}} e^{\frac{log^2(y)}{2}} dy\\
&= e^2
\end{aligned}
$$

$$
\begin{aligned}
Var(Y) &= \mathbf{E}(Y^2) - [\mathbf{E}(Y)]^2\\
&= e^2 - [e^{\frac{1}{2}}]^2\\
&= e^2 - e
\end{aligned}
$$

1.d

$$
\begin{aligned}
Var(Y) &= \mathbf{E}(Y^2) - [\mathbf{E}(Y)]^2\\
&= \mathbf{E}(\mathbf{E}[Y^2|X]) - [\mathbf{E}(\mathbf{E}[Y|X])]^2\\
&= \mathbf{E}(Var(Y|X) + [\mathbf{E}(Y|X)]^2) - [\mathbf{E}(\mathbf{E}[Y|X])]^2\\
&= \mathbf{E}(Var(Y|X)) + \mathbf{E}([\mathbf{E}(Y|X)]^2) - [\mathbf{E}(\mathbf{E}[Y|X])]^2\\
&= \mathbf{E}(Var(Y|X)) + Var(\mathbf{E}(Y|X))
\end{aligned}
$$

2 Basic Regression

$$H = X(X^T X)^{-1} X^T \tag{8}$$
$$\hat{\beta} = (X^T X)^{-1} X^T \mathbf{y} \tag{9}$$
$$\tag{10}$$

2.a

$$\begin{aligned} H\mathbf{y} &= (X(X^T X)^{-1} X^T)\mathbf{y} \\ &= X((X^T X)^{-1} X^T \mathbf{y}) \\ &= X\hat{\beta} \end{aligned}$$

2.b

$$\begin{aligned} HX &= X(X^T X)^{-1} X^T X \\ &= XI \\ &= X \end{aligned}$$

2.c

$$\begin{aligned} H^T &= (X(X^T X)^{-1} X^T)^T \\ &= X^{T^T} (X^T X)^{-1^T} X^T \\ &= X(X^T X^{T^T})^{-1} X^T \\ &= X(X^T X)^{-1} X^T) \\ &= H \end{aligned}$$

2.d

$$\begin{aligned} H^2 &= (X(X^T X)^{-1} X^T)^2 \\ &= (X(X^T X)^{-1} X^T)(X(X^T X)^{-1} X^T) \\ &= X(X^T X)^{-1} I X^T \\ &= (X(X^T X)^{-1} X^T)^2 \\ &= H \end{aligned}$$

2.e

2.f

$$
\begin{aligned}
trace(H) &= trace(X(X^TX)^{-1}X^T) \\
&= trace((X^TX)^{-1}X^TX) \\
&= trace(I_d) \\
&= d = rank(X) \quad \because \text{H is a projection}
\end{aligned}
$$

3 SVD

$$
\begin{aligned}
rank(X) &= r && (11) \\
X &= U\Sigma V^T && (12) \\
\Sigma_{ii} &= \sigma_i U_i = \mathbf{u}_i && (13) \\
V_i &= \mathbf{v}_i && (14)
\end{aligned}
$$

3.a

$$
\begin{aligned}
XX^T\mathbf{u}_i &= U\Sigma V^T(U\Sigma V^T)^T\mathbf{u}_i \\
&= U\Sigma V^T V\Sigma^T U^T\mathbf{u}_i \\
&= U\Sigma\Sigma^T U^T\mathbf{u}_i \\
&= U\Sigma^2 U^T\mathbf{u}_i \\
&= \lambda\mathbf{u}_i \because \text{U is orthogonal } \exists\lambda \\
\lambda &= \sigma_i^2 \because \text{U is unitary } \exists\lambda
\end{aligned}
$$

$$
\begin{aligned}
X^TX\mathbf{v}_i &= (U\Sigma V^T)^T U\Sigma V^T\mathbf{v}_i \\
&= V\Sigma^T U^T U\Sigma V^T\mathbf{v}_i \\
&= V\Sigma\Sigma^T V^T\mathbf{v}_i \\
&= V\Sigma^2 V^T\mathbf{v}_i \\
&= \lambda\mathbf{v}_i \because \text{V is orthogonal } \exists\lambda \\
\lambda &= \sigma_i^2 \because \text{U is unitary } \exists\lambda
\end{aligned}
$$

3.b

$$X\mathbf{u}_i = U\Sigma V^T\mathbf{u}_i$$

3.c

3.d

3.e

$$\begin{aligned}
H &= X(X^TX)^{-1}X^T \\
&= U\Sigma V^T((U\Sigma V^T)^TU\Sigma V^T)^{-1}(U\Sigma V^T)^T \\
&= U\Sigma V^T(V\Sigma^TU^TU\Sigma V^T)^{-1}V\Sigma^TU^T \\
&= U\Sigma V^T(V\Sigma^2V^T)^{-1}V\Sigma^TU^T \\
&= U\Sigma V^TV\Sigma^{-2}V^TV\Sigma^TU^T \\
&= U\Sigma\Sigma^{-2}\Sigma^TU^T \\
&= UU^T
\end{aligned}$$