This week we introduced probability, reviewing basic concepts such as expectation, variance and their properties. We slo covered CDFs and PDFs. Random variables are decribed with cumulative density functions and their probability for a given value is specified with the derivative of the cdf, the probability density function.

A big concept we covered was the weak law of large numbers (WLLN). The WLLN allows us explain that a set of random numbers can converge as the set grows. Intuitively this means that if I have a large set of random numbers, I can appropriately determine the probability of an event.

We also covered Gaussian distributions and their properties, particularly the fact that a random variable equal to a linear combination of random variables characterized by a Gaussian pdf is also a Gaussian pdf.
Gaussian distributions are often seen in robotics problems, because of these properties that make them easier to work with.

Akash Patel
MFML HMWK 7

2. Suppose that two random variables (X, Y) have joint pdf $f_{X,Y}(x,y)$. Find an expression for the pdf $f_Z(z)$ where $Z = X + Y$.

$$f_Z(z) = \frac{\partial F_Z(u)}{\partial u}\Big|_{u=z} \qquad F_Z(u) = P(Z \le u) = \int_{-\infty}^{u} f_Z(z)\,dz$$

$$\frac{dz}{dy} = 1 \qquad dz = dy$$

$$F_Z(u) = \int_{-\infty}^{\infty} F_Z(u|X=\beta)\, f_{X_1}(\beta)\, d\beta = \int_{-\infty}^{\infty}\int_{-\infty}^{u} f_Z(z|X=\beta)\, f_{X_1}(\beta)\, dz\, d\beta$$

$$F_Z(z|X=\beta)$$
$$\int_{-\infty}^{u} f_Z(\beta+y|x=\beta)$$
$$\int_{-\infty}^{u-\beta} f_Y(y|x=\beta)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{u-\beta} f_Y(y|X=\beta)\, f_{X_1}(\beta)\, dy\, d\beta$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{u-\beta} f_{X,Y}(\beta, y)\, dy\, d\beta \quad \longleftarrow \quad f_Y(y|X=\beta) f_X(X=\beta) = f_{XY}(X=\beta, Y=y)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{s} f_{X,Y}(\beta, y)\, dy\, d\beta$$

Let $u - \beta = s$
$\Rightarrow u = s + \beta$
$\Rightarrow \frac{\partial u}{\partial s} = 1$
$\partial u = \partial s$

$$f_Z(z) = \frac{\partial F_Z(u)}{\partial u}\Big|_{u=z}$$
$$= \frac{\partial}{\partial u}\int_{-\infty}^{\infty}\int_{-\infty}^{s} f_{X,Y}(\beta,y)\,dy\,d\beta\Big|_{u=z}$$
$$= \frac{\partial}{\partial s}\int_{-\infty}^{\infty}\int_{-\infty}^{s} f_{X,Y}(\beta,y)\,dy\,d\beta\Big|_{u=z}$$
$$= \int_{-\infty}^{\infty}\frac{\partial}{\partial s}\int_{-\infty}^{s} f_{X,Y}(\beta,y)\,dy\,d\beta\Big|_{u=z}$$

Fundamental Theorem of Calculus

$$= \int_{-\infty}^{\infty} f_{X,Y}(\beta, s)\, d\beta$$

$$\boxed{f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(\beta, u-\beta)\, d\beta} = \int_{-\infty}^{\infty} f_{X,Y}(\beta, z-\beta)\, d\beta$$

How does your expression simplify if X and Y are independent?

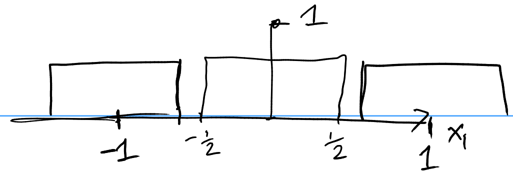
$f_{X,Y}(\beta, z-\beta) = f_X(\beta)\cdot f_Y(z-\beta)$ if X, Y are independent.

$$f_Z(z) = \underbrace{\int_{-\infty}^{\infty} f_X(\beta)\cdot f_Y(z-\beta)\, d\beta}_{\text{a convolution operation.}}$$

3. Let $X_1, X_2, \ldots$ be independent uniform random variables,
$X_n \sim \text{Uniform}(-\frac{1}{2}, \frac{1}{2})$ meaning $f_X(x) = \begin{cases} 1 & -\frac{1}{2} \le x \le \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$

a) What is the density function for $Y = X_1 + X_2 + X_3$.
$f_Y(y) = ?$
Let $X_1 + X_2 = A$
and $A + X_3 = Y$
$Y = A + X_3$
From Problem 2 results:
$f_Y(y) = \int_{-\infty}^{\infty} f_A(a)\, f_{X_3}(y-a)\, da = f_A(y) * f_{X_3}(y)$
$f_A(a) = \int_{-\infty}^{\infty} f_{X_1}(x_1)\, f_{X_2}(a - x_1)\, dx_1 = f_{x_1}(a) * f_{x_2}(a)$
$\Rightarrow f_Y(y) = f_{X_1}(y) * f_{X_2}(y) * f_{X_3}(y)$

First $\underline{f_{x_1}(t) * f_{x_2}(t)}$



$\int_{-\infty}^{\infty} f_{x_1}(x_1) f_{x_2}(t-x_1)\, dx_1$

Case 1) $-\frac{1}{2} \le t + \frac{1}{2} \le \frac{1}{2}$

$-1 \le t \le 0$

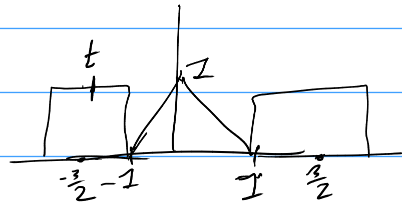
$\int_{-1}^{t} f_{x_1}(x_1) f_{x_2}(t-x_1)\, dx_1 = \int_{-1}^{t} 1\, dx_1 = x_1 \Big|_{-1}^{t} = t+1$

Case 2) $-\frac{1}{2} \le t - \frac{1}{2} \le \frac{1}{2}$

$0 \le t \le 1$

$\int_{t}^{1} 1 dx_1 = x_1 \Big|_{t}^{1} = 1-t$

Now let us convolve $f_{x_3}(t) * (f_{x_1}(t) * f_{x_2}(t))$:



Case 1) $-1 \le t + \frac{1}{2} \le 0 \qquad -\frac{3}{2} \le t \le -\frac{1}{2}$

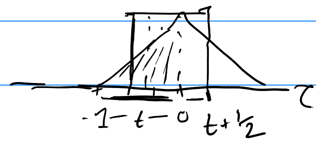
$\int_{-1}^{t+\frac{1}{2}} x_1 + 1\, dx_1 = \frac{x_1^2}{2} + x_1 \Big|_{-1}^{t+1/2} = \frac{(t+\frac{1}{2})^2}{2} + t + \frac{1}{2} - \frac{1}{2} + 1$

$= \frac{t^2}{2} + \frac{(\frac{1}{2})^2}{2} + \frac{2\frac{1}{2}t}{2} + t + 1$

$= \frac{t^2}{2} + \frac{1}{8} + \frac{3t}{2} + 1 = \frac{t^2}{2} + \frac{3t}{2} + \frac{9}{8}$

$= \frac{1}{2}(t^2 + 3t + \frac{9}{4}) = \frac{1}{2}(t + \frac{3}{2})^2$ ✓

Case 2) $0 \le t + \frac{1}{2} \le 1 \qquad -\frac{1}{2} \le t \le \frac{1}{2}$



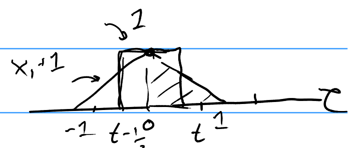
$\int_{t-\frac{1}{2}}^{0} x_1 + 1\, dx_1 + \int_{0}^{t+\frac{1}{2}} 1 - x_1\, dx_1$

$= \frac{x_1^2}{2} + x_1 \Big|_{t-\frac{1}{2}}^{0} + \left(x_1 - \frac{x_1^2}{2}\right)\Big|_{0}^{t+\frac{1}{2}} =$

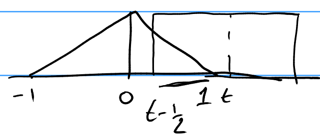
use wolfram

$= \frac{1}{8}(-4t^2 - 4t + 3) + \frac{1}{8}(-4t^2 + 4t + 3)$

$= \frac{1}{8}(-8t^2 + 6) = -t^2 + \frac{3}{4}$ ✓



Case 3) $1 \ge t - \frac{1}{2} \ge 0 \qquad \frac{3}{2} \ge t \ge \frac{1}{2} \qquad \frac{1}{2} \le t \le \frac{3}{2}$



$\int_{t-\frac{1}{2}}^{1} -x_1 + 1\, dx_1 = -\frac{x_1^2}{2} + x_1 \Big|_{t-\frac{1}{2}}^{1} = -\frac{1}{2} + 1 + \frac{(t-\frac{1}{2})^2}{2} - (t - \frac{1}{2})$

$= \frac{1}{2} - t + \frac{1}{2} + \left(\frac{t^2}{2} + \frac{(\frac{1}{2})^2}{2} - \frac{2(\frac{1}{2}t)}{2}\right) = 1 - t + \frac{t^2}{2} + \frac{1}{8} - \frac{1}{2}t$

$= \frac{9}{8} + \frac{t^2}{2} - \frac{3t}{2} = \frac{1}{2}(t^2 - 3t + \frac{9}{4}) = \frac{1}{2}(t - \frac{3}{2})^2$ ✓

Case 4) Outside of $-\frac{3}{2} \le t \le -\frac{3}{2}$ the two functions do not overlap and thus $\underline{f_Y(t) = 0.}$

The B-spline!

Thus $f_y(t) = \begin{cases} \frac{1}{2}\left(t+\frac{3}{2}\right)^2 & -3/2 \le t \le -\frac{1}{2} \\ -t^2+\frac{3}{4} & -\frac{1}{2} \le t \le \frac{1}{2} \\ \frac{1}{2}\left(t-\frac{3}{2}\right)^2 & \frac{1}{2} \le t \le \frac{3}{2} \\ 0 & |t| \ge 3/2 \end{cases}$

b) The moment generating function of a random variable is

$$\phi_X(t) = E[e^{tX}].$$

It is a fact that if $\phi_X(t) = \phi_W(t)$ for all t, then X and W have the same distribution. It is a fact that if $G \sim \text{Normal}(0, \sigma^2)$, then $\phi_G(t) = e^{\sigma^2 t^2/2}$.

Let $Y_N = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} X_n$

Find an expression for $\phi_{Y_N}(t)$.

$$\phi_{Y_N}(t) = E[e^{tY_N}] = E\left[e^{t/\sqrt{N}\left(\sum_{n=1}^{N} X_n\right)}\right]$$

independence of X_n

$$= \prod_{n=1}^{N} E\left[e^{t/\sqrt{N}\, X_n}\right]$$

$$= \prod_{n=1}^{N} \int_{-\infty}^{\infty} e^{\frac{tx_n}{\sqrt{N}}} f_{X_n}(x_n)\, dx_n$$

$$= \prod_{n=1}^{N} \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{\frac{tx_n}{\sqrt{N}}}\, dx_n$$

$$= \prod_{n=1}^{N} \frac{\sqrt{N}}{t} e^{\frac{tx_n}{\sqrt{N}}} \Big|_{-\frac{1}{2}}^{\frac{1}{2}}$$

$$= \prod_{n=1}^{N} \frac{\sqrt{N}}{t}\left(e^{\frac{t}{2\sqrt{N}}} - e^{\frac{t}{-2\sqrt{N}}}\right)$$

$$= \left[\frac{\sqrt{N}}{t}\left(e^{\frac{t}{2\sqrt{N}}} - e^{\frac{t}{-2\sqrt{N}}}\right)\right]^N$$

c) It is a fact that if $\phi(z)$ is a monotonically increasing function, then for any random variable Z,

$$P(Z > u) = P(\phi(Z) > \phi(u))$$

Use $\phi(z) = e^{tz}$ and the Markov inequality to derive a bound on $P(Z_N > u)$, where $$Z_N = \frac{1}{N}\sum_{n=1}^{N} X_n$$

$E[e^{tZ_N}]$ is derived similar to 3b) but we now have $\frac{1}{N}$ instead of $\frac{1}{\sqrt{N}}$.

Markov Inequality: $P(X \geq a) \leq \frac{E[X]}{a}$ for all $a > 0$.

$\Rightarrow P(X > a) < \frac{E[X]}{a}$ for all $a > 0$

$$P(Z_N > u) = P(e^{tZ_N} > e^{tu}) < \frac{E[e^{tZ_N}]}{e^{tu}} = \underline{\left[\frac{N}{t}\left(e^{\frac{t}{2N}} - e^{-\frac{t}{2N}}\right)\right]^N \Big/ e^{tu}}$$

For the special case of $t = 4u/N$, compare this bound, as a function of u, to that obtained using the Chebyshev inequality.

Chebyshev inequality: $P(|X - \mu| > c) \leq \frac{\sigma^2}{c^2}$ for all $c > 0$.

$$P(Z_N > u) = \frac{1}{2}P(|Z_N| > u) = \frac{1}{2}P(|Z_N - 0| > u) = \frac{1}{2}P(|Z_N - \mu| > u) \leq \frac{1}{2}\frac{\sigma^2}{u^2}$$

$$\leq \frac{1}{2}\frac{1}{12Nu^2} = \boxed{\frac{1}{24Nu^2}}$$

$$E[Z_N] = E\left[\frac{1}{N}\sum X_i\right] = \frac{1}{N}\sum E[X_i] = 0 = \mu$$

$$\sigma^2 = Var[Z_N] = Var\left[\frac{1}{N}\sum X_i\right] = \frac{1}{N^2}Var\left[\sum X_i\right] = \frac{1}{N^2}N\,Var(X) \quad X_i \text{ are independent}$$

$$= \frac{Var(X)}{N} = \frac{1}{12N}$$

For $t = 4u/N$:

We see that the Chebyshev inequality gives us a smaller, i.e. tighter upper bound on $P(Z > u)$ for almost all values of u. However, at values really close to 0, the Chebyshev bound blows up whereas the Markov bound does not. Thus the Chebyshev does not give us anything meaningful in that region, since $P(Z > u)$ must be always ≤ 1.

4. $Z = w_1 X_1 + w_2 X_2$ $\qquad Z - w_1X_1 = w_2 X_2$

$$= \begin{bmatrix}1\\1\end{bmatrix}^T \underbrace{\begin{bmatrix} w_1X_1 \\ w_2X_2\end{bmatrix}}_{\Omega} = \begin{bmatrix}1\\1\end{bmatrix}^T \begin{bmatrix} w_1X_1 \\ Z - w_1X_1\end{bmatrix}$$

$$E[Z] = w_1 E[X_1] + w_2 E[X_2] = 0$$

$$R = \begin{bmatrix} Var(w_1X_1) & Cov(w_1X_1, w_2X_2) \\ Cov(w_1X_1, w_2X_2) & Var(w_1X_1)\end{bmatrix}$$

We can now use the following:

$$f_\Omega(w) = \frac{1}{2\pi |R|^{1/2}} \exp\left(-\frac{1}{2} w^T R^{-1} w\right)$$

Let $R^{-1} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22}\end{bmatrix}$

$$f_Z(u) = \int f_\Omega(w)\, d\beta$$
$$= \int f_\Omega\left(\begin{bmatrix}\beta \\ u-\beta\end{bmatrix}\right) d\beta$$
$$= \int \frac{1}{2\pi|R|^{1/2}} \exp\left(-\frac{1}{2}\begin{bmatrix}\beta \\ u-\beta\end{bmatrix}^T \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22}\end{bmatrix}\begin{bmatrix}\beta \\ u-\beta\end{bmatrix}\right)$$

Let us first simplify and complete the square inside the exp.

$$\begin{bmatrix}\beta \\ \alpha\end{bmatrix}^T \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22}\end{bmatrix}\begin{bmatrix}\beta \\ \alpha\end{bmatrix} \qquad \alpha = u - \beta$$

$$= \begin{bmatrix}\beta \\ \alpha\end{bmatrix}^T \begin{bmatrix} \beta r_{11} + \alpha r_{12} \\ \beta r_{21} + \alpha r_{22}\end{bmatrix} = r_{11}\beta^2 + \alpha\beta r_{12} + \alpha\beta r_{21} + \alpha^2 r_{22}$$

$$= r_{11}\beta^2 + (r_{12} + r_{21})\alpha\beta + \alpha^2 r_{22} \qquad \alpha = u - \beta$$
$$= r_{11}\beta^2 + (r_{12}+r_{21})(u-\beta)\beta + (u-\beta)^2 r_{22}$$
$$= r_{11}\beta^2 + (r_{12}+r_{21})(u\beta - \beta^2) + (u^2 + \beta^2 - 2u\beta) r_{22}$$
$$= r_{11}\beta^2 + (r_{12}+r_{21})u\beta - (r_{12}+r_{21})\beta^2 + r_{22}u^2 + r_{22}\beta^2 - 2u\beta r_{22}$$

$a = r_{11} - r_{12} - r_{21} + r_{22}$

$$= (r_{11} - r_{12} - r_{21} + r_{22})\beta^2 + r_{22}u^2 + (r_{12} + r_{21} - 2r_{22})u\beta$$

$b = (r_{12} + r_{21} - 2r_{22})u$

$c = r_{22}u^2$

$b = b'u$

$c = c'u^2$

$$= a\beta^2 + b\beta + c \qquad \text{complete the square}$$
$$= a\left(\beta + \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a}$$

expanding back out: $a\left(\beta + \frac{b'u}{2a}\right)^2 + c'u^2 - \frac{b'^2u^2}{4a}$

$$= a\left(\beta + \frac{b'u}{2a}\right)^2 + \frac{4ac' - b'^2}{4a} u^2$$

Plug back into pdf

$$f_Z(u) = \int_{-\infty}^{\infty} \frac{1}{2\pi|R|^{1/2}} \exp\left(-\frac{1}{2}\left[a\left(\beta + \frac{b'u}{2a}\right)^2 + \frac{4ac'-b'^2}{4a}u^2\right]\right) d\beta$$

$$= \frac{1}{2\pi|R|^{1/2}} \exp\left(-\frac{1}{2}\frac{4ac'-b'^2}{4a}u^2\right)\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}a\left(\beta + \frac{b'u}{2a}\right)^2\right) d\beta$$

$$= \frac{1}{2\pi|R|^{1/2}} \exp\left(-\frac{1}{2}\frac{4ac'-b'^2}{4a}u^2\right)\underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi a^{-1}}} \exp\left(-\frac{1}{2}\frac{\left(\beta + \frac{b'u}{2a}\right)^2}{a^{-1}}\right) d\beta}_{\text{Gaussian w.r.t } \beta \,=\, 1} \cdot \sqrt{2\pi a^{-1}}$$

$$= \frac{\sqrt{2\pi a^{-1}}}{2\pi|R|^{1/2}} \exp\left(-\frac{1}{2}\frac{u^2}{\frac{4a}{4ac'-b'^2}}\right) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{a|R|}} \exp\left(-\frac{1}{2}u^2 \Big/ \left(\frac{4a}{4ac'-b'^2}\right)\right)$$

$$|R| = \left|\begin{bmatrix} var(w_1X_1) & cov(w_1X_1, w_2X_2) \\ cov(w_1X_1, w_2X_2) & var(w_2X_2) \end{bmatrix}\right| = \left|\begin{bmatrix} w_1^2 Var(X_1) & E[w_1w_2X_1X_2] \\ E[w_1w_2X_1X_2] & w_2^2 var(X_2) \end{bmatrix}\right| = \left|\begin{bmatrix} w_1^2\sigma_1^2 & w_1w_2\gamma \\ w_1w_2\gamma & w_2^2\sigma_2^2 \end{bmatrix}\right|$$

$$= w_1^2\sigma_1^2 w_2^2\sigma_2^2 - w_1^2w_2^2\gamma^2$$

$$R^{-1} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \frac{1}{w_1^2\sigma_1^2w_2^2\sigma_2^2 - w_1^2w_2^2\gamma^2} \begin{bmatrix} w_2^2\sigma_2^2 & -w_1w_2\gamma \\ -w_1w_2\gamma & w_1^2w_1^2 \end{bmatrix}$$

$$r_{11} = \frac{w_2^2\sigma_2^2}{|R|} \qquad a = r_{11} - r_{12} - r_{21} + r_{22} \qquad c' = r_{22}$$

$$r_{21} = r_{12} = \frac{-w_1w_2\gamma}{|R|} \qquad = (w_2^2\sigma_2^2 + w_1^2\sigma_1^2 + 2w_1w_2\gamma)/|R| \qquad b' = r_{12} + r_{21} - 2r_{22}$$

$$r_{22} = \frac{w_1^2\sigma_1^2}{|R|} \qquad a|R| = w_2^2\sigma_2^2 + w_1^2\sigma_1^2 + 2w_1w_2\gamma$$

$$\frac{4\cancel{a}}{4ac' - b'^2} = a\left(\frac{4}{4ac'-b'^2}\right) = a\left(\frac{4}{4(r_{11}-r_{12}-r_{21}+r_{22})(r_{22}) - (r_{12}+r_{21}-2r_{22})^2}\right)$$

$$= a\frac{4}{4(r_{11}+r_{22}-2r_{12})(r_{22}) - (2r_{12}-2r_{22})^2} = a\frac{1}{r_{11}r_{22} + \cancel{r_{22}^2} - 2r_{12}\cancel{r_{22}} - r_{12}^2 - \cancel{r_{22}^2} + 2\cancel{r_{12}r_{22}}}$$

$$= a\frac{1}{r_{11}r_{22} - r_{12}^2} = a\frac{|R|^2}{w_2^2\sigma_2^2w_1^2\sigma_1^2 - w_1^2w_2^2\gamma^2} = a\frac{|R|^2}{|R|} = a|R| \checkmark$$

Simplifying expression:

$$f_2(u) = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{a|R|}} \exp\left(-\frac{1}{2}u^2 / \left(\frac{4a}{4ac'-b'^2}\right)\right)$$

$$= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{a|R|}} \exp\left(-\frac{1}{2}u^2/a|R|\right)$$ ← This is a Gaussian distribution with mean $= 0$ and variance $= a|R| = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\gamma$

$R = Var(X) = E[XX^T]$

5. Let $X \in \mathbb{R}^D$ be a Gaussian random vector, $X \sim Normal(\mu, R)$

a) Using your answer to the previous question, show that for $w \in \mathbb{R}^D$, $Y = w^T X$ is a scalar Gaussian random variable with mean $w^T\mu$ and variance $w^T R w$.

$Y = w^T X$ can be written as $Y = \sum_d^D w_d X_d$

we have alredy shown that $Z = \begin{bmatrix}1\\1\end{bmatrix}^T \begin{bmatrix}w_1 x_1\\ w_2 x_2\end{bmatrix} = w^T X$ is a scalar gaussian from proben 4.

$E[Y] = E[\sum_d^D w_d X_d] = \sum_d^D E[w_d X_d] = \sum_d^D w_d E[X_d] = \sum_d^D w_d \mu_d$

$= w^T \mu$ ✓

$Var[Y]$ $\qquad E[(w^T X)^T{}^2] = E[w^T X X^T w] =$

$= Var[w^T X] = E[(w^T X)^2] = E[w^T X X^T w] = w^T \underbrace{E[XX^T]}_{R} w = w^T R w$ ✓

b) Show that if $R_{i,j} = 0$, then X_i and X_j are independent, that is
$f_{X_i X_j}(x_i, x_j) = f_{X_i}(x_i) f_{X_j}(x_j)$

Let us take the case with $X = \begin{bmatrix} X_i \\ X_j \end{bmatrix}$ then

$f_{x_i,x_j}(x_i,x_j) = f_X(x) = \frac{1}{2\pi |R|^{1/2}} \exp\left(-\frac{1}{2}(X-\mu)^T R^{-1} (X-\mu)\right)$

$R = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} r_{11} & 0 \\ 0 & r_{22} \end{bmatrix}$

$R^{-1} = \frac{1}{r_{11} r_{22}} \begin{bmatrix} r_{22} & 0 \\ 0 & r_{11} \end{bmatrix} = \begin{bmatrix} \frac{1}{r_{11}} & 0 \\ 0 & \frac{1}{r_{22}} \end{bmatrix}$

$|R| = r_{11} r_{22}$

r_{11} is $var(X_i)$
r_{12} is $var(X_j)$

$\begin{bmatrix} x_i - \mu_i \\ x_j - \mu_j \end{bmatrix} \begin{bmatrix} \frac{1}{r_{11}} & 0 \\ 0 & \frac{1}{r_{22}} \end{bmatrix} \begin{bmatrix} x_i - \mu_i \\ x_j - \mu_j \end{bmatrix}$

$= (x_i - \mu_i)^2 \frac{1}{r_{11}} + (x_j - \mu_j) \frac{1}{r_{22}}$

$f_X(x) = \frac{1}{2\pi \sqrt{r_{11} r_{12}}} \exp\left(-\frac{1}{2}\frac{1}{r_{11}}(x_i - \mu_i)^2 + -\frac{1}{2}\frac{1}{r_{22}}(x_j - \mu_j)\right)$

$= \frac{1}{2\pi\sqrt{r_{11}}} \exp\left(-\frac{1}{2}\frac{1}{r_{11}}(x_i - \mu_i)^2\right) \cdot \frac{1}{2\pi\sqrt{r_{22}}} \exp\left(-\frac{1}{2}\frac{1}{r_{22}}(x_j - \mu_j)^2\right)$

$= f_{X_i}(x_i) \cdot f_{X_j}(x_j)$ ✓

6. Suppose that random vector $X \in \mathbb{R}^D$ is distributed $X \sim \text{Normal}(0, I)$. Describe how to choose a $D \times D$ matrix A so that $Y = AX$ aboves $Y \sim \text{Normal}(0, R)$ for an arbitrary covariance matrix R.

We can write $Y = \sum_{d}^{D} a_d^T X = \sum_{d}^{D} Y_d$ where a_d is the d^{th} row vector of A.

From problem 5) we know that $Y_d \sim N(0, a_d^T I a)$

and that $Y \sim N(0, AIA^T) \Rightarrow Y \sim N(0, AA^T)$

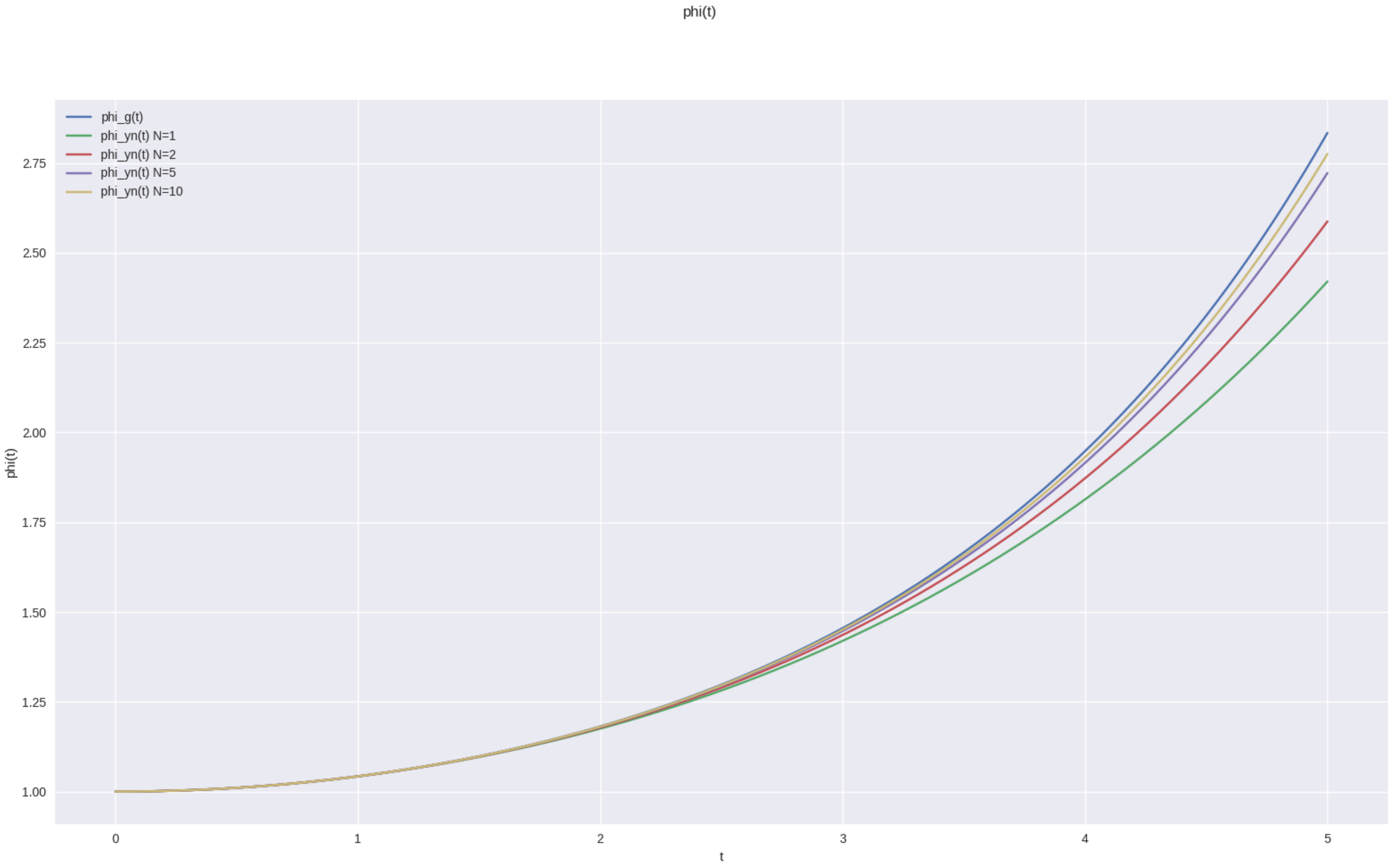
Let us now equate AA^T with R.

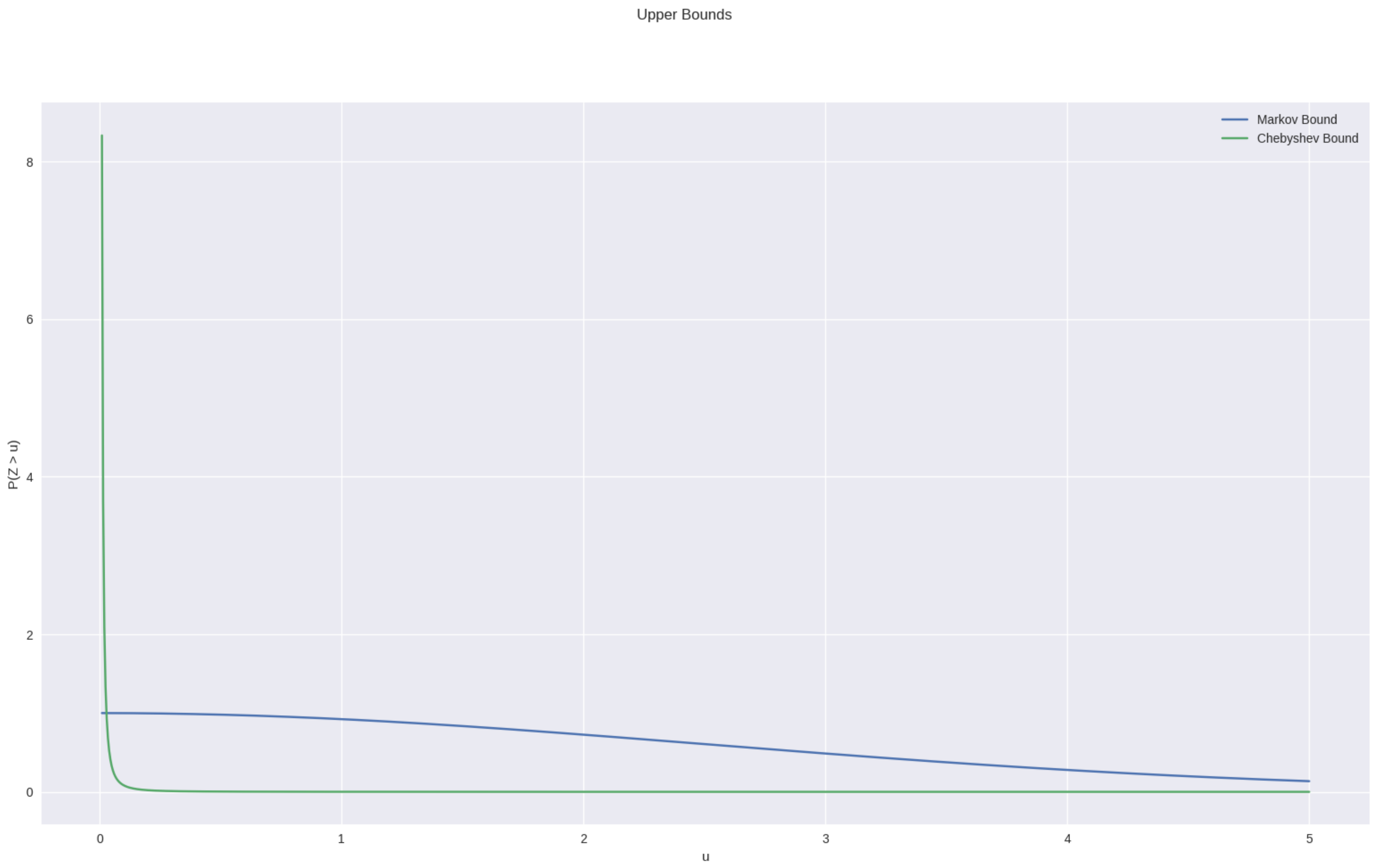
$$AA^T = R \xrightarrow[\text{evd since } R \text{ is sym+def}]{} AA^T = V \Lambda V^T \Rightarrow AA^T = \underbrace{V\Lambda^{1/2}}\underbrace{\Lambda^{T/2} V^T}$$

Thus $\boxed{A = V\Lambda^{1/2}}$

```python
"""Problem 3."""
import matplotlib as mpl
import matplotlib.pyplot as plt

import numpy as np

mpl.style.use('seaborn')


def phi_yn(t_val, n_val):
    """phi_yn."""
    return (np.sqrt(n_val)/t_val *
            (np.exp(t_val/(2 * np.sqrt(n_val))) -
             np.exp(t_val/(-2 * np.sqrt(n_val)))))**n_val


def phi_g(t_val, sigma):
    """phi_g."""
    return np.exp(sigma**2 * t_val ** 2/2)


def part_b():
    """Part b."""
    print('Part b')

    sigma_sqrd = 1/12
    sigma = np.sqrt(sigma_sqrd)
    n_list = [1, 2, 5, 10]
    t_vec = np.linspace(0.001, 5, 1000)

    fig = plt.figure()
    fig.suptitle('phi(t)')
    axes = fig.add_subplot(111)

    axes.plot(t_vec, phi_g(t_vec, sigma), label='phi_g(t)')

    for n_val in n_list:
        axes.plot(t_vec, phi_yn(t_vec, n_val),
                  label='phi_yn(t) N=' + str(n_val))

    axes.set_xlabel('t')
    axes.set_ylabel('phi(t)')
    axes.legend()

    plt.show()


part_b()


def phi_zn(t_val, n_val):
    """phi_zn."""
    return (n_val/t_val *
            (np.exp(t_val/(2 * n_val)) -
             np.exp(t_val/(-2 * n_val))))**n_val


def markov_bound(t_val, u_val, n_val):
    """Markov bound."""
    return phi_zn(t_val, n_val)/np.exp(t_val * u_val)


def chebyshev_bound(u_val, n_val):
    """Chebyshev bound."""
```

```
    return 1 / (24 * n_val * u_val**2)


def part_c():
    """Part c."""
    print('Part c')

    fig = plt.figure()
    fig.suptitle('Upper Bounds')
    axes = fig.add_subplot(111)

    n_val = 50
    u_vec = np.linspace(0.01, 5, 1000)

    t_vec = 4 * u_vec / n_val

    axes.plot(u_vec, markov_bound(t_vec, u_vec, n_val), label='Markov Bound')
    axes.plot(u_vec, chebyshev_bound(u_vec, n_val), label='Chebyshev Bound')

    axes.set_xlabel('u')
    axes.set_ylabel('P(Z > u)')
    axes.legend()

    plt.show()


part_c()
```


phi(t)
phi_g(t)
phi_yn(t) N=1
phi_yn(t) N=2
phi_yn(t) N=5
phi_yn(t) N=10
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
phi(t)
0
1
2
3
4
5
t


Upper Bounds
Markov Bound
Chebyshev Bound
P(Z > u)
u
0
2
4
6
8
0
1
2
3
4
5