Multivariate Gaussians

Sinho Chewi

Spring 2017

University of California, Berkeley

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0, 1)$.

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0, 1)$. Density?

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0,1)$. Density?

$$f_Z(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_i^2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} z^\top z\right).$$

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0,1)$. Density?

$$f_Z(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z_i^2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}z^\top z\right).$$

$A \in \mathbb{R}^{n\times n}$ is an invertible matrix.

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0,1)$. Density?

$$f_Z(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_i^2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} z^\top z\right).$$

$A \in \mathbb{R}^{n \times n}$ is an invertible matrix.

$\mu \in \mathbb{R}^n$ is a vector of means.

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0,1)$. Density?

$$f_Z(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_i^2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} z^\top z\right).$$

$A \in \mathbb{R}^{n \times n}$ is an invertible matrix.

$\mu \in \mathbb{R}^n$ is a vector of means.

Change of variables: $X = AZ + \mu$.

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0,1)$. Density?

$$f_Z(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_i^2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} z^\top z\right).$$

$A \in \mathbb{R}^{n \times n}$ is an invertible matrix.

$\mu \in \mathbb{R}^n$ is a vector of means.

Change of variables: $X = AZ + \mu$.

$$f_X(x)\,\mathrm{d}x = f_Z(z)\,\mathrm{d}z.$$

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0,1)$. Density?

$$f_Z(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_i^2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} z^\top z\right).$$

$A \in \mathbb{R}^{n\times n}$ is an invertible matrix.

$\mu \in \mathbb{R}^n$ is a vector of means.

Change of variables: $X = AZ + \mu$.

$$f_X(x)\,\mathrm{d}x = f_Z(z)\,\mathrm{d}z.$$

Multivariate calculus: $\mathrm{d}x = \det(A)\,\mathrm{d}z$.

Multivariate Gaussian Density

$Z_1, \ldots, Z_n$ are i.i.d. $\mathcal{N}(0,1)$. Density?

$$f_Z(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} z_i^2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} z^\top z\right).$$

$A \in \mathbb{R}^{n \times n}$ is an invertible matrix.

$\mu \in \mathbb{R}^n$ is a vector of means.

Change of variables: $X = AZ + \mu$.

$$f_X(x)\,\mathrm{d}x = f_Z(z)\,\mathrm{d}z.$$

Multivariate calculus: $\mathrm{d}x = \det(A)\,\mathrm{d}z$.

$$\begin{aligned} f_X(x) &= \frac{1}{\det(A)} f_Z(A^{-1}(x-\mu)) \\ &= \frac{1}{(2\pi)^{n/2}\det(A)} \exp\left(-\frac{1}{2}(x-\mu)^\top (A^{-1})^\top A^{-1}(x-\mu)\right). \end{aligned}$$

Multivariate Gaussian Density

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\det(A)} \exp\left(-\frac{1}{2}(x-\mu)^\top (AA^\top)^{-1}(x-\mu)\right).$$

Multivariate Gaussian Density

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\det(A)} \exp\left(-\frac{1}{2}(x-\mu)^\top (AA^\top)^{-1}(x-\mu)\right).$$

Let $\Sigma = AA^\top$.

Multivariate Gaussian Density

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\det(A)} \exp\left(-\frac{1}{2}(x-\mu)^\top (AA^\top)^{-1}(x-\mu)\right).$$

Let $\Sigma = AA^\top$.

$\det(A) = \det(\Sigma)^{1/2}$.

Multivariate Gaussian Density

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\det(A)} \exp\left(-\frac{1}{2}(x-\mu)^\top (AA^\top)^{-1}(x-\mu)\right).$$

Let $\Sigma = AA^\top$.

$\det(A) = \det(\Sigma)^{1/2}$.

$$f_X(x) = \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right).$$

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\det(A)} \exp\left(-\frac{1}{2}(x-\mu)^\top (AA^\top)^{-1}(x-\mu)\right).$$

Let $\Sigma = AA^\top$.
$\det(A) = \det(\Sigma)^{1/2}$.

$$f_X(x) = \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right).$$

We assumed A is invertible (non-degenerate case).

Multivariate Gaussian Density

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\det(A)} \exp\left(-\frac{1}{2}(x-\mu)^\top (AA^\top)^{-1}(x-\mu)\right).$$

Let $\Sigma = AA^\top$.
$\det(A) = \det(\Sigma)^{1/2}$.

$$f_X(x) = \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right).$$

We assumed A is invertible (non-degenerate case).
Σ is the covariance matrix: $\Sigma_{i,j} = \operatorname{cov}(X_i, X_j)$.

Multivariate Gaussian Density

$$f_X(x) = \frac{1}{(2\pi)^{n/2}\det(A)} \exp\left(-\frac{1}{2}(x-\mu)^\top (AA^\top)^{-1}(x-\mu)\right).$$

Let $\Sigma = AA^\top$.

$\det(A) = \det(\Sigma)^{1/2}$.

$$f_X(x) = \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right).$$

We assumed A is invertible (non-degenerate case).

Σ is the covariance matrix: $\Sigma_{i,j} = \mathrm{cov}(X_i, X_j)$.

Proof: $\mathrm{cov}((AZ)_i, (AZ)_j) = \mathrm{cov}(\sum_{k=1}^n A_{i,k}Z_k, \sum_{l=1}^n A_{j,l}Z_l)$

$= \sum_{k=1}^n A_{i,k}A_{j,k} = (AA^\top)_{i,j}$.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Then, $BX = (BA)Z + B\mu$.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Then, $BX = (BA)Z + B\mu$.

$\implies$ BX is jointly Gaussian.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Then, $BX = (BA)Z + B\mu$.

$\implies$ BX is jointly Gaussian.

Conversely, if any linear combination of $X_1, \dots, X_n$ is Gaussian, then $X_1, \dots, X_n$ are jointly Gaussian.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Then, $BX = (BA)Z + B\mu$.

$\implies$ BX is jointly Gaussian.

Conversely, if any linear combination of $X_1, \ldots, X_n$ is Gaussian, then $X_1, \ldots, X_n$ are jointly Gaussian.

Example: Let $X \sim \mathcal{N}(0, 1)$ and $\Pr(Z = \pm 1) = 1/2$ be independent.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Then, $BX = (BA)Z + B\mu$.

$\implies$ BX is jointly Gaussian.

Conversely, if any linear combination of $X_1, \ldots, X_n$ is Gaussian, then $X_1, \ldots, X_n$ are jointly Gaussian.

Example: Let $X \sim \mathcal{N}(0, 1)$ and $\Pr(Z = \pm 1) = 1/2$ be independent.

X and $Y = XZ$ are marginally Gaussian.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Then, $BX = (BA)Z + B\mu$.

$\implies$ BX is jointly Gaussian.

Conversely, if any linear combination of $X_1, \ldots, X_n$ is Gaussian, then $X_1, \ldots, X_n$ are jointly Gaussian.

Example: Let $X \sim \mathcal{N}(0, 1)$ and $\Pr(Z = \pm 1) = 1/2$ be independent.

X and $Y = XZ$ are marginally Gaussian.

X and $Y = XZ$ are not jointly Gaussian.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Then, $BX = (BA)Z + B\mu$.

$\implies$ BX is jointly Gaussian.

Conversely, if any linear combination of $X_1, \dots, X_n$ is Gaussian, then $X_1, \dots, X_n$ are jointly Gaussian.

Example: Let $X \sim \mathcal{N}(0, 1)$ and $\Pr(Z = \pm 1) = 1/2$ be independent.

X and $Y = XZ$ are marginally Gaussian.

X and $Y = XZ$ are not jointly Gaussian.

$X - Y = (1 - Z)X$.

Linear Combinations of Joint Gaussians

Linear combinations of jointly Gaussian random variables are jointly Gaussian.

Write $X = AZ + \mu$.

Then, $BX = (BA)Z + B\mu$.

$\implies$ BX is jointly Gaussian.

Conversely, if any linear combination of $X_1, \ldots, X_n$ is Gaussian, then $X_1, \ldots, X_n$ are jointly Gaussian.

Example: Let $X \sim \mathcal{N}(0,1)$ and $\Pr(Z = \pm 1) = 1/2$ be independent.

X and $Y = XZ$ are marginally Gaussian.

X and $Y = XZ$ are not jointly Gaussian.

$X - Y = (1 - Z)X$.

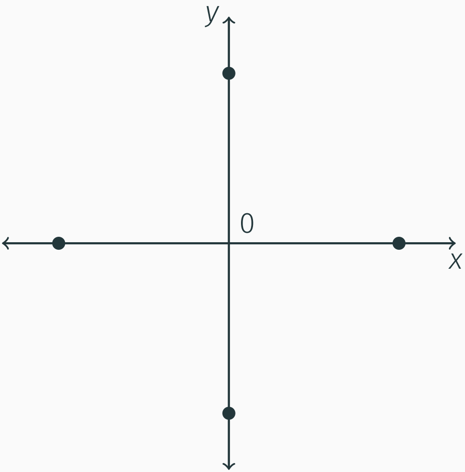
$\Pr(X - Y = 0) = 1/2$.

Uncorrelated Random Variables

Uncorrelated random variables are not necessarily independent.

Uncorrelated Random Variables

Uncorrelated random variables are not necessarily independent.



Uncorrelated Gaussians

If $X_1, \ldots, X_n$ are uncorrelated, jointly Gaussian, they are independent.

Uncorrelated Gaussians

If $X_1, \ldots, X_n$ are uncorrelated, jointly Gaussian, they are independent.

uncorrelated $\implies \Sigma$ is diagonal

Uncorrelated Gaussians

If $X_1, \ldots, X_n$ are uncorrelated, jointly Gaussian, they are independent.

uncorrelated $\implies \Sigma$ is diagonal

Let $\Sigma^{-1} = \text{diag}(\sigma_1^{-2}, \ldots, \sigma_n^{-2})$.

Uncorrelated Gaussians

If $X_1, \ldots, X_n$ are uncorrelated, jointly Gaussian, they are independent.

uncorrelated $\implies$ Σ is diagonal

Let $\Sigma^{-1} = \mathrm{diag}(\sigma_1^{-2}, \ldots, \sigma_n^{-2})$.

$$f_X(x) = \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right)$$

Uncorrelated Gaussians

If $X_1, \ldots, X_n$ are uncorrelated, jointly Gaussian, they are independent.

uncorrelated $\implies$ Σ is diagonal

Let $\Sigma^{-1} = \operatorname{diag}(\sigma_1^{-2}, \ldots, \sigma_n^{-2})$.

$$\begin{aligned} f_X(x) &= \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right) \\ &= \frac{1}{(2\pi)^{n/2}\sigma_1 \cdots \sigma_n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{1}{\sigma_i^2}(x_i-\mu_i)^2\right) \end{aligned}$$

Uncorrelated Gaussians

If $X_1, \ldots, X_n$ are uncorrelated, jointly Gaussian, they are independent.

uncorrelated $\implies$ Σ is diagonal

Let $\Sigma^{-1} = \mathrm{diag}(\sigma_1^{-2}, \ldots, \sigma_n^{-2})$.

$$
\begin{aligned}
f_X(x) &= \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right) \\
&= \frac{1}{(2\pi)^{n/2}\sigma_1 \cdots \sigma_n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{1}{\sigma_i^2}(x_i-\mu_i)^2\right) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2}(x_i-\mu_i)^2\right).
\end{aligned}
$$

Uncorrelated Gaussians

If $X_1, \ldots, X_n$ are uncorrelated, jointly Gaussian, they are independent.

uncorrelated $\implies$ Σ is diagonal

Let $\Sigma^{-1} = \operatorname{diag}(\sigma_1^{-2}, \ldots, \sigma_n^{-2})$.

$$\begin{aligned}
f_X(x) &= \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right) \\
&= \frac{1}{(2\pi)^{n/2}\sigma_1 \cdots \sigma_n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{1}{\sigma_i^2}(x_i-\mu_i)^2\right) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2}(x_i-\mu_i)^2\right).
\end{aligned}$$

Uncorrelated Gaussians

If $X_1, \ldots, X_n$ are uncorrelated, jointly Gaussian, they are independent.

uncorrelated $\implies$ Σ is diagonal

Let $\Sigma^{-1} = \operatorname{diag}(\sigma_1^{-2}, \ldots, \sigma_n^{-2})$.

$$
\begin{aligned}
f_X(x) &= \frac{1}{(2\pi)^{n/2}(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\right) \\
&= \frac{1}{(2\pi)^{n/2}\sigma_1 \cdots \sigma_n} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \frac{1}{\sigma_i^2}(x_i-\mu_i)^2\right) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{1}{2\sigma_i^2}(x_i-\mu_i)^2\right).
\end{aligned}
$$

Joint density factors into product of marginal densities.

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Proof:

$$\Pr(X \in A, \phi(Y) \in B) = \Pr(X \in A, Y \in \phi^{-1}(B))$$

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Proof:

$$\begin{aligned}\Pr(X \in A, \phi(Y) \in B) &= \Pr(X \in A, Y \in \phi^{-1}(B)) \\ &= \Pr(X \in A) \Pr(Y \in \phi^{-1}(B))\end{aligned}$$

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Proof:

$$\begin{aligned}
\Pr(X \in A, \phi(Y) \in B) &= \Pr(X \in A, Y \in \phi^{-1}(B)) \\
&= \Pr(X \in A) \Pr(Y \in \phi^{-1}(B)) \\
&= \Pr(X \in A) \Pr(\phi(Y) \in B).
\end{aligned}$$

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Proof:

$$\begin{aligned}
\Pr(X \in A, \phi(Y) \in B) &= \Pr(X \in A, Y \in \phi^{-1}(B)) \\
&= \Pr(X \in A) \Pr(Y \in \phi^{-1}(B)) \\
&= \Pr(X \in A) \Pr(\phi(Y) \in B).
\end{aligned}$$

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Proof:

$$\begin{aligned}\Pr(X \in A, \phi(Y) \in B) &= \Pr(X \in A, Y \in \phi^{-1}(B)) \\ &= \Pr(X \in A)\Pr(Y \in \phi^{-1}(B)) \\ &= \Pr(X \in A)\Pr(\phi(Y) \in B).\end{aligned}$$

Also recall:

If X and Y are independent, zero-mean, they are orthogonal.

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Proof:

$$\begin{aligned}
\Pr(X \in A, \phi(Y) \in B) &= \Pr(X \in A, Y \in \phi^{-1}(B)) \\
&= \Pr(X \in A) \Pr(Y \in \phi^{-1}(B)) \\
&= \Pr(X \in A) \Pr(\phi(Y) \in B).
\end{aligned}$$

Also recall:

If X and Y are independent, zero-mean, they are orthogonal.

$Y - L[Y \mid X]$ is orthogonal to any linear function of X.

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Proof:

$$\begin{aligned}\Pr(X \in A, \phi(Y) \in B) &= \Pr(X \in A, Y \in \phi^{-1}(B)) \\ &= \Pr(X \in A) \Pr(Y \in \phi^{-1}(B)) \\ &= \Pr(X \in A) \Pr(\phi(Y) \in B).\end{aligned}$$

Also recall:

If X and Y are independent, zero-mean, they are orthogonal.

$Y - L[Y \mid X]$ is orthogonal to any linear function of X.

$Y - E[Y \mid X]$ is orthogonal to any function of x.

Lemma (Independent Random Variables)

If X and Y are independent, then so are X and $\phi(Y)$.

Proof:

$$\begin{aligned}
\Pr(X \in A, \phi(Y) \in B) &= \Pr(X \in A, Y \in \phi^{-1}(B)) \\
&= \Pr(X \in A)\Pr(Y \in \phi^{-1}(B)) \\
&= \Pr(X \in A)\Pr(\phi(Y) \in B).
\end{aligned}$$

Also recall:

If X and Y are independent, zero-mean, they are orthogonal.

$Y - L[Y \mid X]$ is orthogonal to any linear function of X.

$Y - E[Y \mid X]$ is orthogonal to any function of x.

For jointly Gaussian X and Y, $L[Y \mid X]$ is Gaussian.

Jointly Gaussian LLSE $=$ MMSE

For jointly Gaussian X, Y, the LLSE and the MMSE coincide.

Jointly Gaussian LLSE = MMSE

For jointly Gaussian X, Y, the LLSE and the MMSE coincide.

$Y - L[Y \mid X]$ and X are orthogonal.

Jointly Gaussian LLSE $=$ MMSE

For jointly Gaussian X, Y, the LLSE and the MMSE coincide.

$Y - L[Y \mid X]$ and X are orthogonal.

$Y - L[Y \mid X]$ and X are uncorrelated.

Jointly Gaussian LLSE $=$ MMSE

For jointly Gaussian X, Y, the LLSE and the MMSE coincide.

$Y - L[Y \mid X]$ and X are orthogonal.

$Y - L[Y \mid X]$ and X are uncorrelated.

$Y - L[Y \mid X]$ and X are independent.

Jointly Gaussian LLSE $=$ MMSE

For jointly Gaussian X, Y, the LLSE and the MMSE coincide.

$Y - L[Y \mid X]$ and X are orthogonal.

$Y - L[Y \mid X]$ and X are uncorrelated.

$Y - L[Y \mid X]$ and X are independent.

$Y - L[Y \mid X]$ and $\phi(X)$ are independent.

Jointly Gaussian LLSE = MMSE

For jointly Gaussian X, Y, the LLSE and the MMSE coincide.

$Y - L[Y \mid X]$ and X are orthogonal.

$Y - L[Y \mid X]$ and X are uncorrelated.

$Y - L[Y \mid X]$ and X are independent.

$Y - L[Y \mid X]$ and $\phi(X)$ are independent.

$Y - L[Y \mid X]$ and $\phi(X)$ are orthogonal.

Jointly Gaussian LLSE = MMSE

For jointly Gaussian X, Y, the LLSE and the MMSE coincide.

$Y - L[Y \mid X]$ and X are orthogonal.

$Y - L[Y \mid X]$ and X are uncorrelated.

$Y - L[Y \mid X]$ and X are independent.

$Y - L[Y \mid X]$ and $\phi(X)$ are independent.

$Y - L[Y \mid X]$ and $\phi(X)$ are orthogonal.

$\implies Y - L[Y \mid X]$ is orthogonal to any $\phi(X)$, so

Jointly Gaussian LLSE $=$ MMSE

For jointly Gaussian X, Y, the LLSE and the MMSE coincide.

$Y - L[Y \mid X]$ and X are orthogonal.

$Y - L[Y \mid X]$ and X are uncorrelated.

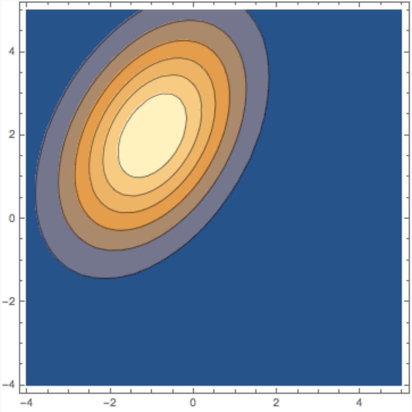
$Y - L[Y \mid X]$ and X are independent.

$Y - L[Y \mid X]$ and $\phi(X)$ are independent.

$Y - L[Y \mid X]$ and $\phi(X)$ are orthogonal.

$\implies$ $Y - L[Y \mid X]$ is orthogonal to any $\phi(X)$, so $L[Y \mid X] = E[Y \mid X]$.

$$X \sim \mathcal{N}\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\right).$$



Eigenvectors of $\boldsymbol{\Sigma}$:

$$\lambda_1 = 3.62, \qquad v_1 = \begin{bmatrix} 0.526 \\ 0.851 \end{bmatrix}$$

$$\lambda_2 = 1.38, \qquad v_2 = \begin{bmatrix} -0.851 \\ 0.525 \end{bmatrix}$$