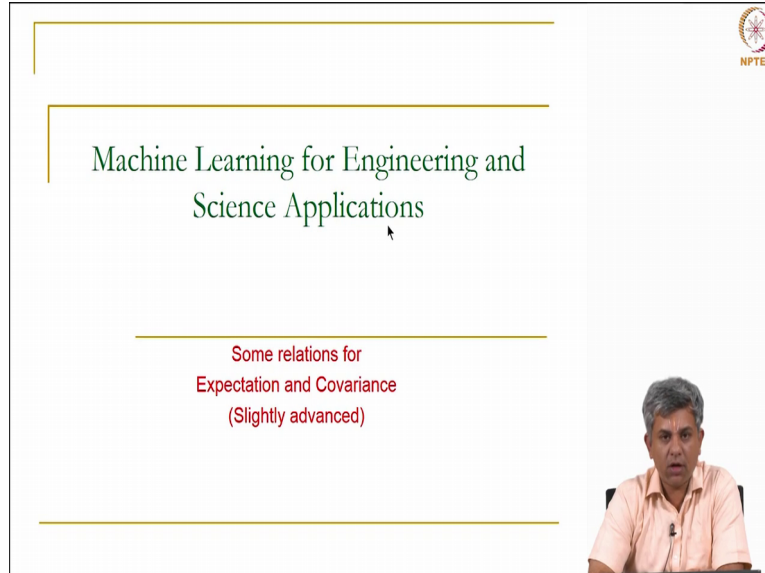


Machine Learning for Engineering and Science Applications
Professor Dr. Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology Madras
Some Relations for Expectation and Covariance (Slightly Advanced)

(Refer Slide Time: 00:14)

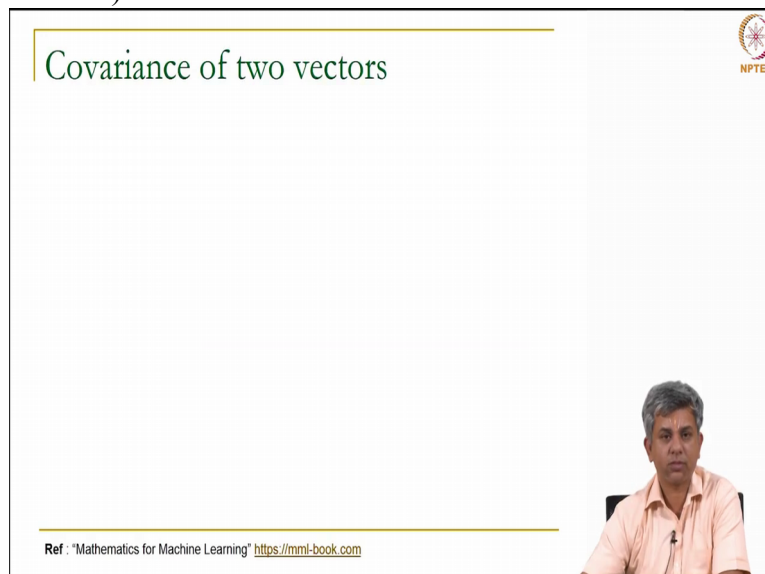


Machine Learning for Engineering and Science Applications

Some relations for
Expectation and Covariance
(Slightly advanced)

In this video we will be looking at some more relations for expectation and covariance. These are slightly advanced relations. We will be using these only rarely in the course. So in case you do not understand this portion that is Ok, you will get to understand it a little better as the course progresses.

(Refer Slide Time: 00:33)



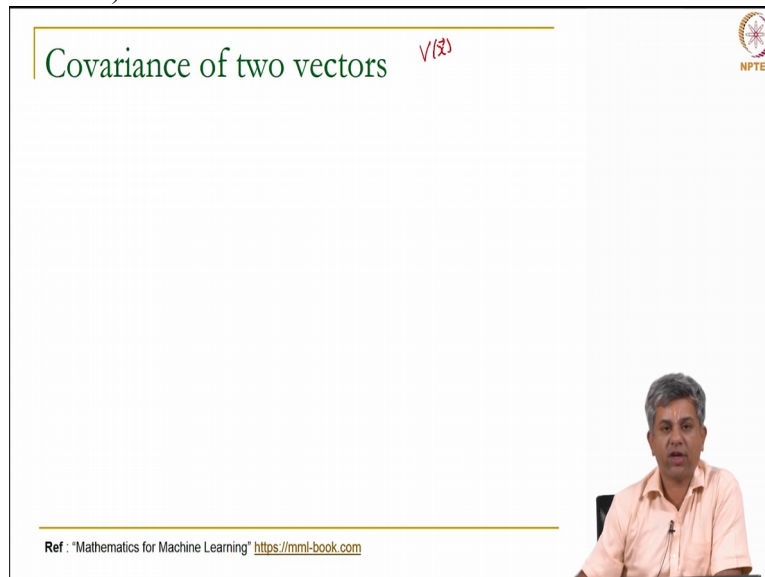
Covariance of two vectors

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

So please do not panic in case it looks a little bit unfamiliar to you.

So we have looked at the variance for a single vector, variance of x vector.

(Refer Slide Time: 00:46)



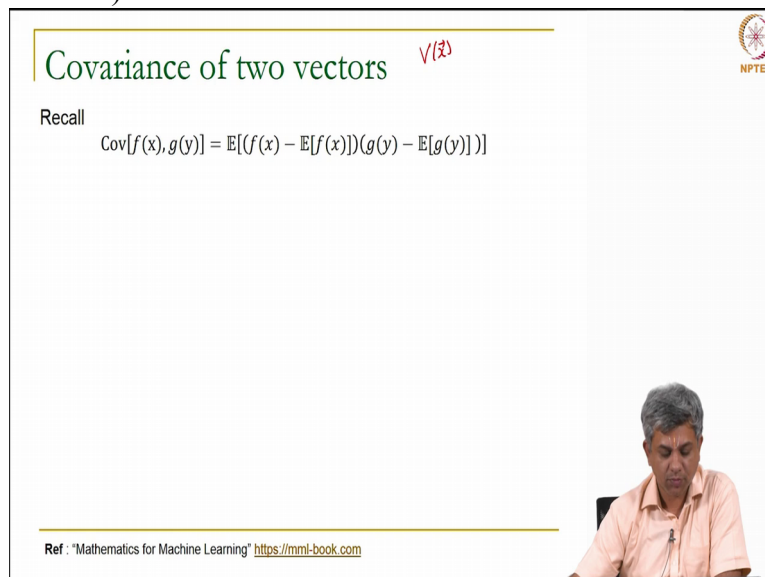
Covariance of two vectors $V(x)$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

The slide features a title "Covariance of two vectors" in green, a red handwritten note "V(x)" next to it, and the NPTEL logo in the top right corner. A small video inset of a man in a light orange shirt is visible in the bottom right corner of the slide frame.

Now we are going to look at the covariance of two vectors.

(Refer Slide Time: 00:50)



Covariance of two vectors $V(x)$

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

The slide features a title "Covariance of two vectors" in green, a red handwritten note "V(x)" next to it, and the NPTEL logo in the top right corner. Below the title, the word "Recall" is followed by the covariance formula. A small video inset of a man in a light orange shirt is visible in the bottom right corner of the slide frame.

So remember that for scalar functions f of x and g of y we had defined the covariance as the deviation from the expectation of f of x multiplied by the deviation from the expectation of g of y and expectation



(Refer Slide Time: 01:08)

Covariance of two vectors ^{V(x)}

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$
$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>



of the whole thing, Ok.

We also, you might also remember that the simpler definition was that of single x and y and that



(Refer Slide Time: 01:16)

Covariance of two vectors ^{V(x)}

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$
$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$
$$\text{Cov}[\mathbf{x}, \mathbf{x}]_{i,j} = \text{Cov}[x_i, x_j]$$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>



definition is given here. Finally we had looked at covariance of two vectors, Ok. This can be called as the variance. Remember just like covariance of x x is called

(Refer Slide Time: 01:33)


Covariance of two vectors $\sqrt{\bar{x}}$

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$
$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

$\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ $\text{Cov}(x,x) = \mathbb{V}[x]$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>



variance of x which is the square of the standard deviation, Ok.

Now we had defined something called the covariance matrix.

(Refer Slide Time: 01:52)

Covariance of two vectors $\sqrt{\bar{x}}$


Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$
$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

$\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ $\text{Cov}(x,x) = \mathbb{V}[x]$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x

Ref: "Mathematics for Machine Learning" <https://mml-book.com>



This was a matrix of all covariances, you can recollect this from the previous video. This is simply a matrix of covariance of, let us say the first element will be covariance of x_1 and x_1 , so on

(Refer Slide Time: 02:09)

Covariance of two vectors V(x)

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix

$$\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j] \quad \left[\begin{array}{l} \text{Cov}(x,x) = \mathbb{V}[x] \\ \text{Block A} \end{array} \right]$$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x

Ref.: "Mathematics for Machine Learning" <https://mml-book.com>

and so forth. Remember x is a vector and x_i is the i th component of the vector x .

Now let us say we denote by μ , this is standard notation, mean or expectation is denoted by μ . Let us say μ is the expectation of the vector x and since it is a vector, μ is also going to be a vector. Again if you look at the previous videos you will see that the first element of μ is the first element

(Refer Slide Time: 02:37)

Covariance of two vectors V(x)

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix

$$\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j] \quad \left[\begin{array}{l} \text{Cov}(x,x) = \mathbb{V}[x] \\ \text{Block A} \end{array} \right] \quad \mu_1 = \mathbb{E}[x_1]$$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x


Ref.: "Mathematics for Machine Learning" <https://mml-book.com>

of the expectation so on and so forth. That is a full vector.

So let us say μ is the expectation of the random vector

(Refer Slide Time: 02:44)

Covariance of two vectors V(x)



Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$


$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix \rightarrow $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ [Root(A)] $\text{Cov}(x,x) = \sqrt{[x]}$ $\mu_1 = \mathbb{E}[x]$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x

Then,

$$\text{Cov}[x, x] = \mathbb{E}[(x - \mu)(x - \mu)^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$




Ref: "Mathematics for Machine Learning" <https://mml-book.com>

x then covariance of x comma x which we have seen before, Ok can be written as x minus μ times x minus μ transpose. This is the way that we had defined it before, Ok.

So covariance of x simply by extending this is x minus μ . Remember now, x is a vector, x minus μ is also a vector

(Refer Slide Time: 03:09)

Covariance of two vectors V(x)



Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$


$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix \rightarrow $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ [Root(A)] $\text{Cov}(x,x) = \sqrt{[x]}$ $\mu_1 = \mathbb{E}[x]$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x

Then,

$$\text{Cov}[x, x] = \mathbb{E}[(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$




Ref: "Mathematics for Machine Learning" <https://mml-book.com>

and you are taking a transpose, Ok. Why are we taking a transpose? What is the size of this? This is a n cross 1

(Refer Slide Time: 03:21)

Covariance of two vectors √(x)



Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$


$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix → $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ [Root(A)] $\text{Cov}(x,x) = \sqrt{[x]}$
 $\mu_1 = \mathbb{E}[x]$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x

Then,

$$\text{Cov}[x, x] = \mathbb{E}[(\overset{[n \times 1]}{\bar{x}} - \overset{[1 \times n]}{\bar{\mu}})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$




Ref: "Mathematics for Machine Learning" <https://mml-book.com>

vector. Suppose I take a transpose,

(Refer Slide Time: 03:25)

Covariance of two vectors √(x)



Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$


$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix → $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ [Root(A)] $\text{Cov}(x,x) = \sqrt{[x]}$
 $\mu_1 = \mathbb{E}[x]$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x

Then,

$$\text{Cov}[x, x] = \mathbb{E}[(\overset{[n \times 1]}{\bar{x}} - \overset{[1 \times n]}{\bar{\mu}})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$



Ref: "Mathematics for Machine Learning" <https://mml-book.com>

this is 1 cross n. Therefore what you will get is I have written m cross m, so this should be m. And this can be m

(Refer Slide Time: 03:33)

Covariance of two vectors V(x)

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix \rightarrow $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ [Root of A] $\text{Cov}(x, x) = \mathbb{V}[x]$ $\mu_1 = \mathbb{E}[x]$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x

Then,

$$\text{Cov}[x, x] = \mathbb{E}[(\bar{x} - \bar{\mu})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

also, Ok. So covariance of x comma x can now be defined in this way. Expectation of x x transpose minus expectation of x multiplied by expectation of x transpose, Ok.

Remember this, if x is a scalar this simply comes to expectation of x square minus expectation of x the whole square.

(Refer Slide Time: 04:00)

Covariance of two vectors V(x)

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix \rightarrow $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ [Root of A] $\text{Cov}(x, x) = \mathbb{V}[x]$ $\mu_1 = \mathbb{E}[x]$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x for scalar x $\mathbb{E}[x^2] - \mathbb{E}[x]^2$

Then,

$$\text{Cov}[x, x] = \mathbb{E}[(\bar{x} - \bar{\mu})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

This we have seen before also. As I

(Refer Slide Time: 04:03)

Covariance of two vectors $\sqrt{\Sigma}$

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix \rightarrow $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ $\left[\begin{smallmatrix} \text{Cov}(x,x) \\ \text{Root(A)} \end{smallmatrix} \right] \mu_1 = \mathbb{E}[x]$

Let $\mu \in \mathbb{R}^m$ be the expectation of the random vector x for each $x \in \mathbb{E}[x] - \mathbb{E}[x]^2$

Then, [m x 1] [1 x m]

$$\text{Cov}[x, x] = \mathbb{E}[(\bar{x} - \bar{\mu})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$

We use the notation $\text{Var}[x] = \text{Cov}[x, x]$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

said before we used the notation that variance of x is equal to covariance of

(Refer Slide Time: 04:07)

Covariance of two vectors $\sqrt{\Sigma}$

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix \rightarrow $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ $\left[\begin{smallmatrix} \text{Cov}(x,x) \\ \text{Root(A)} \end{smallmatrix} \right] \mu_1 = \mathbb{E}[x]$

Let $\mu \in \mathbb{R}^m$ be the expectation of the random vector x for each $x \in \mathbb{E}[x] - \mathbb{E}[x]^2$

Then, [m x 1] [1 x m]

$$\text{Cov}[x, x] = \mathbb{E}[(\bar{x} - \bar{\mu})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$

We use the notation $\text{Var}[x] = \text{Cov}[x, x]$

Similarly, $\text{Cov}[x, y] = \mathbb{E}[xy^T] - \mathbb{E}[x]\mathbb{E}[y]^T = \text{Cov}[y, x]^T \in \mathbb{R}^{m \times n}$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

x comma x.

Now similarly, just very, very similar to this idea, we can now define covariance of x comma y. Now let us say x is in R

(Refer Slide Time: 04:19)

Covariance of two vectors V(x)

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix \rightarrow $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ $\left[\begin{smallmatrix} \text{Roots} \\ \text{of } \lambda \end{smallmatrix} \right]$ $\text{Cov}(x, x) = \sqrt{\text{Var}[x]}$ $\mu_1 = \mathbb{E}[x]$

Let $\bar{\mu} \in \mathbb{R}^n$ be the expectation of the random vector x for each x $\mathbb{E}[x] = \mathbb{E}[x]^T$

Then, $\text{Cov}[x, x] = \mathbb{E}[(\bar{x} - \bar{\mu})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$

We use the notation $\text{Var}[x] = \text{Cov}[x, x]$

Similarly, $\text{Cov}[x, y] = \mathbb{E}[xy^T] - \mathbb{E}[x]\mathbb{E}[y]^T = \text{Cov}[y, x]^T \in \mathbb{R}^{m \times n}$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

m, Ok that is, and y belongs to R n

(Refer Slide Time: 04:23)

Covariance of two vectors V(x)

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \overline{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix \rightarrow $\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j]$ $\left[\begin{smallmatrix} \text{Roots} \\ \text{of } \lambda \end{smallmatrix} \right]$ $\text{Cov}(x, x) = \sqrt{\text{Var}[x]}$ $\mu_1 = \mathbb{E}[x]$

Let $\bar{\mu} \in \mathbb{R}^n$ be the expectation of the random vector x for each x $\mathbb{E}[x] = \mathbb{E}[x]^T$

Then, $\text{Cov}[x, x] = \mathbb{E}[(\bar{x} - \bar{\mu})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$

We use the notation $\text{Var}[x] = \text{Cov}[x, x]$

Similarly, $\text{Cov}[x, y] = \mathbb{E}[xy^T] - \mathbb{E}[x]\mathbb{E}[y]^T = \text{Cov}[y, x]^T \in \mathbb{R}^{m \times n}$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

which means that x is now a m cross 1 vector, y is a n cross 1 vector. So y transpose is now

(Refer Slide Time: 04:34)

Covariance of two vectors V(x)

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \bar{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix

$$\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j] \quad \left[\begin{array}{l} \text{Root(A)} \\ \text{Cov}(x,x) = \sqrt{\text{Var}[x]} \\ \mu_1 = \mathbb{E}[x] \end{array} \right]$$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x for each x $\mathbb{E}[x] = \mathbb{E}[x]^2$

Then,

$$\text{Cov}[x, x] = \mathbb{E}[(\bar{x} - \bar{\mu})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$

We use the notation $\text{Var}[x] = \text{Cov}[x, x]$

Similarly, $\text{Cov}[x, y] = \mathbb{E}[xy^T] - \mathbb{E}[x]\mathbb{E}[y]^T = \text{Cov}[y, x]^T \in \mathbb{R}^{m \times n}$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

1 cross n, Ok.

So we would write instead of $x x^T$, we would write expectation of $x y^T$ minus expectation of x multiplied by expectation of y^T , this is simply a definition. This is just a way we define covariance of x comma y .

You can show easily that covariance of x comma y is the same as covariance of y comma x the whole transpose. I would suggest that you try this as an exercise.

(Refer Slide Time: 05:06)

Covariance of two vectors V(x)

Recall

$$\text{Cov}[f(x), g(y)] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])(g(y) - \mathbb{E}[g(y)])]$$

$$\text{Cov}[x, y] = \bar{xy} - \bar{x}\bar{y} = \mathbb{E}[xy] - \mathbb{E}[x]\mathbb{E}[y]$$

Covariance Matrix

$$\text{Cov}[x, x]_{i,j} = \text{Cov}[x_i, x_j] \quad \left[\begin{array}{l} \text{Root(A)} \\ \text{Cov}(x,x) = \sqrt{\text{Var}[x]} \\ \mu_1 = \mathbb{E}[x] \end{array} \right]$$

Let $\mu \in \mathbb{R}^n$ be the expectation of the random vector x for each x $\mathbb{E}[x] = \mathbb{E}[x]^2$

Then,

$$\text{Cov}[x, x] = \mathbb{E}[(\bar{x} - \bar{\mu})(\bar{x} - \bar{\mu})^T] = \mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T \in \mathbb{R}^{m \times m}$$

We use the notation $\text{Var}[x] = \text{Cov}[x, x]$ Exercise

Similarly, $\text{Cov}[x, y] = \mathbb{E}[xy^T] - \mathbb{E}[x]\mathbb{E}[y]^T = \text{Cov}[y, x]^T \in \mathbb{R}^{m \times n}$

Ref: "Mathematics for Machine Learning" <https://mml-book.com>

Now another idea that we are going

(Refer Slide Time: 05:14)

Sums of random variables

- Recall that expectation is a linear operator

to look at is the sums of two random variables. Remember that we had

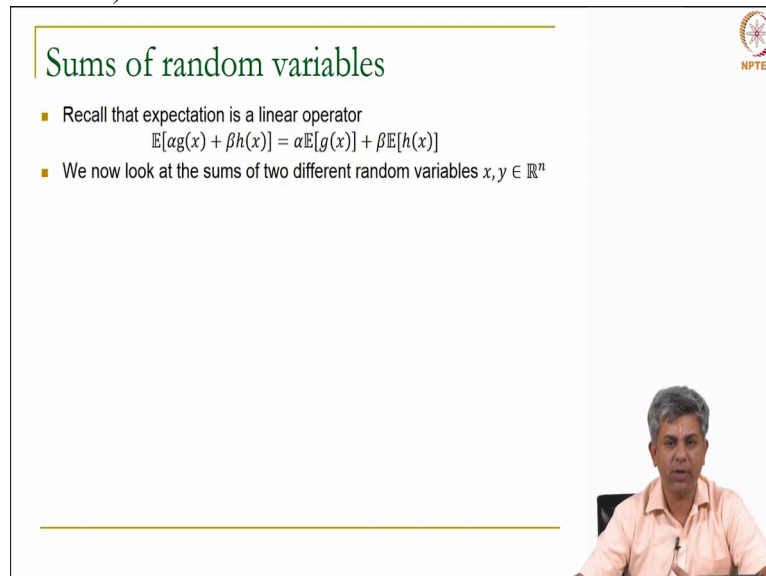
(Refer Slide Time: 05:18)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$$

already discussed that expectation is a linear operator so that if you have expectation of alpha g plus beta h, it is going to be alpha times expectation g

(Refer Slide Time: 05:27)

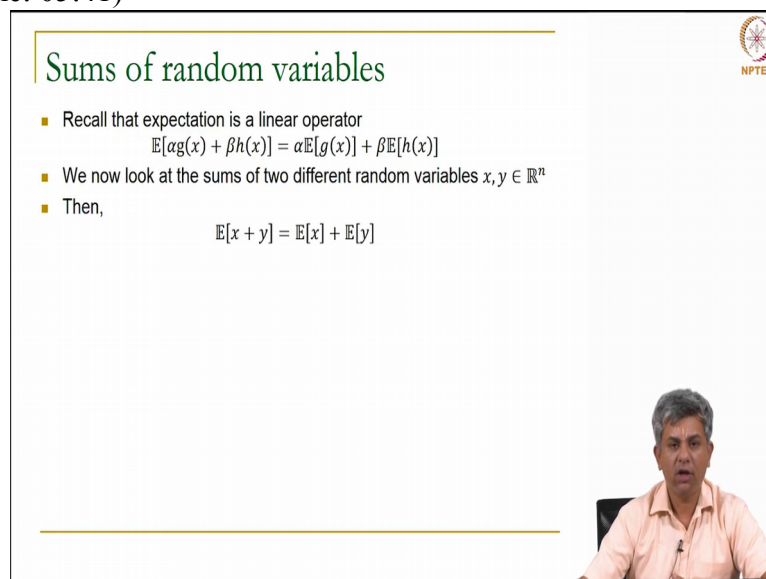


The slide is titled "Sums of random variables" in green text. It contains two bullet points: "Recall that expectation is a linear operator" followed by the equation $\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$, and "We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$ ". The NPTEL logo is in the top right corner. A speaker is visible in the bottom right corner of the slide frame.

plus beta times expectation h.

Now we are going to extend this idea to that of two random variables, two different random variables let us say x and y both are now of the same size n cross 1 ,

(Refer Slide Time: 05:41)





The slide is titled "Sums of random variables" in green text. It contains three bullet points: "Recall that expectation is a linear operator" followed by the equation $\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$, "We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$ ", and "Then," followed by the equation $\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$. The NPTEL logo is in the top right corner. A speaker is visible in the bottom right corner of the slide frame.

then we know that since expectation is linear, expectation of x plus y is simply going to be expectation of x plus expectation of y .

(Refer Slide Time: 05:48)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,
$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$
$$\mathbb{E}[\alpha x + \beta y] = \alpha \mathbb{E}[x] + \beta \mathbb{E}[y]$$





Similarly, expectation of alpha x plus beta y is going to be alpha times expectation x plus beta

(Refer Slide Time: 05:55)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,
$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$
$$\mathbb{E}[\alpha x + \beta y] = \alpha \mathbb{E}[x] + \beta \mathbb{E}[y]$$

Variances are a bit more involved.



times expectation y. Variances however behave a slightly more differently,



(Refer Slide Time: 06:01)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[ag(x) + \beta h(x)] = a\mathbb{E}[g(x)] + \beta\mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,
$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$

$$\mathbb{E}[ax + \beta y] = a\mathbb{E}[x] + \beta\mathbb{E}[y]$$

Variances are a bit more involved.
$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$



Ok. So variation of x plus y , variance of x plus y is not simply variance of x plus variance of y . You have



(Refer Slide Time: 06:11)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[ag(x) + \beta h(x)] = a\mathbb{E}[g(x)] + \beta\mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,
$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$

$$\mathbb{E}[ax + \beta y] = a\mathbb{E}[x] + \beta\mathbb{E}[y]$$

Variances are a bit more involved.
$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$



these two additional terms as well, Ok. You have these covariance terms, covariance of x comma y and covariance of y comma x .



These two are not equal. They are the transposes of

(Refer Slide Time: 06:22)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,
$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$
$$\mathbb{E}[\alpha x + \beta y] = \alpha \mathbb{E}[x] + \beta \mathbb{E}[y]$$

Variances are a bit more involved.

$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$
$$\text{Var}[\alpha x] = \alpha \text{Var}[x]$$




each other. If you take variance of alpha x, it should be alpha square, sorry,

(Refer Slide Time: 06:30)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,
$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$
$$\mathbb{E}[\alpha x + \beta y] = \alpha \mathbb{E}[x] + \beta \mathbb{E}[y]$$

Variances are a bit more involved.

$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$
$$\text{Var}[\alpha x] = \alpha^2 \text{Var}[x]$$


it is alpha square times

(Refer Slide Time: 06:33)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,
$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$



$$\mathbb{E}[\alpha x + \beta y] = \alpha \mathbb{E}[x] + \beta \mathbb{E}[y]$$

Variations are a bit more involved.

$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$

$$\text{Var}[\alpha x] = \alpha^2 \text{Var}[x]$$

Note that, if x, y are independent then $\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y]$



variance of x . Now notice that if x and y are independent, Ok if x and y are independent we had this discussion in an earlier slide, that if x and y are independent then covariance of x y is 0,

(Refer Slide Time: 06:53)

Sums of random variables

- Recall that expectation is a linear operator
$$\mathbb{E}[\alpha g(x) + \beta h(x)] = \alpha \mathbb{E}[g(x)] + \beta \mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,
$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$



$$\mathbb{E}[\alpha x + \beta y] = \alpha \mathbb{E}[x] + \beta \mathbb{E}[y]$$

Variations are a bit more involved.

$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$

$$\text{Var}[\alpha x] = \alpha^2 \text{Var}[x]$$

Note that, if x, y are independent then $\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y]$




Ok.

Similarly, covariance of

(Refer Slide Time: 07:00)

Sums of random variables



- Recall that expectation is a linear operator

$$\mathbb{E}[ag(x) + \beta h(x)] = a\mathbb{E}[g(x)] + \beta\mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,

$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$

$$\mathbb{E}[ax + \beta y] = a\mathbb{E}[x] + \beta\mathbb{E}[y]$$


Variations are a bit more involved.

$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$

$$\text{Var}[ax] = a^2\text{Var}[x]$$

Cov[x, y] = 0 || Cov[y, x] = 0


Note that, if x, y are independent then $\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y]$



$\text{Cov}[x, y]$ is also 0, Ok. So this gives only under the condition that x and y are independent, do we get that variance of x plus y is variance of x plus variance of y , Ok.

(Refer Slide Time: 07:15)

Sums of random variables



- Recall that expectation is a linear operator

$$\mathbb{E}[ag(x) + \beta h(x)] = a\mathbb{E}[g(x)] + \beta\mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,

$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$

$$\mathbb{E}[ax + \beta y] = a\mathbb{E}[x] + \beta\mathbb{E}[y]$$

Variations are a bit more involved.


$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$

$$\text{Var}[ax] = a^2\text{Var}[x]$$

Cov[x, y] = 0 || Cov[y, x] = 0

Note that, if x, y are independent then $\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y]$

④ What is $\text{Var}[x - y]$?




Now as an exercise please think about what happens to variance of x minus y .

So if combine these two relationships, I would again suggest this as an exercise. I will give you the final expression. So variance of x minus y will be variance of x plus variance of y , notice the plus, it is not minus, minus covariance of x y minus covariance of y x . Again I would suggest this as a quick exercise for you to try

(Refer Slide Time: 07:55)

Sums of random variables



- Recall that expectation is a linear operator

$$\mathbb{E}[ag(x) + \beta h(x)] = a\mathbb{E}[g(x)] + \beta\mathbb{E}[h(x)]$$
- We now look at the sums of two different random variables $x, y \in \mathbb{R}^n$
- Then,

$$\mathbb{E}[x + y] = \mathbb{E}[x] + \mathbb{E}[y]$$

$$\mathbb{E}[ax + \beta y] = a\mathbb{E}[x] + \beta\mathbb{E}[y]$$

Variations are a bit more involved.


$$\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y] + \text{Cov}[x, y] + \text{Cov}[y, x]$$

\uparrow
 $\text{Var}[ax] = a^2\text{Var}[x]$
 $\text{Cov}[x, y] = 0 \quad \parallel \quad \text{Cov}[y, x] = 0$

Note that, if x, y are independent then $\text{Var}[x + y] = \text{Var}[x] + \text{Var}[y]$

$\text{Var}[x - y] = \text{Var}[x] + \text{Var}[y] - \text{Cov}[x, y] - \text{Cov}[y, x] = \text{Var}[x] + \text{Var}[y]$

What is $\text{Var}[x - y]$?




out, Ok.


If you ignore the minus parts, minus covariance $x y$ etc, notice that even if you have the difference of two variables, the errors will still add, Ok. Variance is like the error, Ok. It is like the variation.

You would have seen this in simple, you know experimental measurements perhaps a little bit before that if you have two variables, you know I make one length minus another length it does not mean the error subtract, the error still add because the errors go as the variance, Ok. So therefore the errors add there.

(Refer Slide Time: 08:26)

Affine transformations





The final idea that I would like

(Refer Slide Time: 08:29)

Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$

to discuss in this video is that of an affine transformation. An affine transformation is nothing but another name for a linear transformation.

So let us say you have two random variables or two variables x and y . x is a vector, y is a vector. A is a matrix.

(Refer Slide Time: 08:48)

Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$

You can transform, y can be seen as x with a linear transform, b is also a vector here of course in order for our dimensions to match, Ok.

So x is a vector, b is a vector, A is a matrix and y is another vector. So y is Ax plus b . Now if I have this,

(Refer Slide Time: 09:08)

The slide is titled "Affine transformations" in green. It contains two bullet points: "Affine transform : Transformation of variables of the form $y = Ax + b$ " and "Question: Given the mean and variance matrix of x find them for y ". There are handwritten red annotations: "Matrix" above the equation, and arrows pointing from x to A , A to y , and b to y . The NPTEL logo is in the top right corner. A video inset at the bottom right shows a man in a light orange shirt speaking.

suppose I know the mean of x , or suppose I know the variance of x , can I find out the mean and variance of y ? It turns out that

(Refer Slide Time: 09:17)

The slide is titled "Affine transformations" in green. It contains three bullet points: "Affine transform : Transformation of variables of the form $y = Ax + b$ ", "Question: Given the mean and variance matrix of x find them for y ", and "Then, since expectation is a linear operator". There are handwritten red annotations: "Matrix" above the equation, and arrows pointing from x to A , A to y , and b to y . The NPTEL logo is in the top right corner. A video inset at the bottom right shows the same man in a light orange shirt speaking.

there are mathematical relationships. Expectation works out as we expect, Ok.

So

(Refer Slide Time: 09:23)

Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator

$$E_y[y] = E_x[Ax + b] = AE_x[x] + b = A\mu + b$$

expectation of y , we will write the expressions here. Please notice this, the subscript. This expectation is over the variable y . This expectation is over the variable x .

So expectation of y is expectation of $Ax + b$ which since, expectation is linear we simply take out A which is a constant matrix, outside. A times expectation x plus b , μ of course is the notation for expectation of

(Refer Slide Time: 09:55)

Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator

$$E_y[y] = E_x[Ax + b] = AE_x[x] + b = A\mu + b$$

x , Ok.

So this is quite simple as far the expectation is concerned.

(Refer Slide Time: 10:01)

Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$ Matrix
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator $E_x[x]$

$$E_y[y] = E_x[Ax + b] = AE_x[x] + b = A\mu + b$$

$$\begin{aligned} V_y[y] &= V_x[Ax + b] \\ &= V_x[Ax] \\ &= \text{Cov}[Ax, Ax] \\ &= E[(Ax)(Ax)^T] - E[Ax]E[Ax]^T \\ &= E[Axx^T A^T] - AE[x]E[x]^T A^T \\ &= A(E[xx^T] - E[x]E[x]^T)A^T \\ &= ACov[x, x]A^T = A\Sigma A^T \end{aligned}$$

Now variance, we will go through this slowly is once again slightly more involved, Ok. Let us look at it step by step.

So variance of y is, of course y is defined as $Ax + b$ so

(Refer Slide Time: 10:15)

Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$ Matrix
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator $E_x[x]$


$$E_y[y] = E_x[Ax + b] = AE_x[x] + b = A\mu + b$$

$$\begin{aligned} V_y[y] &= V_x[Ax + b] \\ &= V_x[Ax] \\ &= \text{Cov}[Ax, Ax] \\ &= E[(Ax)(Ax)^T] - E[Ax]E[Ax]^T \\ &= E[Axx^T A^T] - AE[x]E[x]^T A^T \\ &= A(E[xx^T] - E[x]E[x]^T)A^T \\ &= ACov[x, x]A^T = A\Sigma A^T \end{aligned}$$

this is fine. Variance of $Ax + b$ is same as variance of x because variance of something x plus a constant is the same as

(Refer Slide Time: 10:34)


Affine transformations



- Affine transform : Transformation of variables of the form $y = Ax + b$ Matrix
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator $E_x[x]$

$$E_y[y] = E_x[Ax + b] = AE_x[x] + b = A\mu + b$$

$$\begin{aligned}
 V_y[y] &= V_x[Ax + b] \\
 &= V_x[Ax] \rightarrow \text{Because } V_x(\text{const} + b) = V_x(x) \\
 &= \text{Cov}[Ax, Ax] \\
 &= E[(Ax)(Ax)^T] - E[Ax]E[Ax]^T \\
 &= E[Axx^T A^T] - AE[x]E[x]^T A^T \\
 &= A(E[xx^T] - E[x]E[x]^T)A^T \\
 &= ACov[x, x]A^T = A\Sigma A^T
 \end{aligned}$$



variance of x .

Why is that? You can prove this very easily mathematically but let us just look at it very simply from a physical point of view.

Variance measures the difference from the mean, Ok or the distance from the mean. If I change the variable by a constant the mean will also go up by a constant and that constant subtracts out.

We are only looking at the difference from the mean and we are not looking at the actual value of the variable, Ok. Since that is the case so whether I add b or not, I am going to get the same variance, Ok.

Now variance of Ax was defined as covariance of the variable with itself, Ax comma Ax . Now we can write it out based on the definition that we had earlier. Remember that covariance of the variable with itself was expectation of yy^T minus expectation of y expectation of y^T . That is what is


(Refer Slide Time: 11:32)


Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator

$$\mathbb{E}_x[y] = \mathbb{E}_x[Ax + b] = A\mathbb{E}_x[x] + b = A\mu + b$$

$$\begin{aligned} \mathbb{V}_y[y] &= \mathbb{V}_x[Ax + b] \\ &= \mathbb{V}_x[Ax] \rightarrow \text{Because } \mathbb{V}_x(x+b) = \mathbb{V}_x(x) \\ &= \text{Cov}[Ax, Ax] \\ &= \mathbb{E}[(Ax)(Ax)^T] - \mathbb{E}[Ax]\mathbb{E}[Ax]^T \\ &= \mathbb{E}[Axx^T A^T] - A\mathbb{E}[x]\mathbb{E}[x]^T A^T \\ &= A(\mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T)A^T \\ &= ACov[x, x]A^T = A\Sigma A^T \end{aligned}$$





used here, Ok.

Instead of y I have Ax . A is A times x transpose, A times expectation of x transpose. Ok. Now we can open this up noting the fact that A times x transpose is the same as


(Refer Slide Time: 11:49)


Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator

$$\mathbb{E}_x[y] = \mathbb{E}_x[Ax + b] = A\mathbb{E}_x[x] + b = A\mu + b$$

$$\begin{aligned} \mathbb{V}_y[y] &= \mathbb{V}_x[Ax + b] \\ &= \mathbb{V}_x[Ax] \rightarrow \text{Because } \mathbb{V}_x(x+b) = \mathbb{V}_x(x) \\ &= \text{Cov}[Ax, Ax] \\ &= \mathbb{E}[(Ax)(Ax)^T] - \mathbb{E}[Ax]\mathbb{E}[Ax]^T \\ &= \mathbb{E}[Axx^T A^T] - A\mathbb{E}[x]\mathbb{E}[x]^T A^T \\ &= A(\mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T)A^T \\ &= ACov[x, x]A^T = A\Sigma A^T \end{aligned}$$





x transpose A transpose.

So that is opened up here. Now also I have used the fact that expectation of Ax is A times expectation of x . Similarly expectation of Ax transpose is the same as expectation of x transpose A transpose which is the same as expectation of x transpose multiplied by A transpose.

(Refer Slide Time: 12:13)

Affine transformations


Matrix

- Affine transform : Transformation of variables of the form $y = Ax + b$
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator $E_x[x]$

$$E_y[y] = E_x[Ax + b] = AE_x[x] + b = A\mu + b$$

$$\begin{aligned} V_y[y] &= V_x[Ax + b] \\ &= V_x[Ax] \rightarrow \text{Because } V_x(x+b) = V(x) \\ &= \text{Cov}[Ax, Ax] \\ &= E[(Ax)(Ax)^T] - E[Ax]E[Ax]^T \\ &= E[Axx^T A^T] - AE[x]E[x]^T A^T \\ &= A(E[xx^T] - E[x]E[x]^T)A^T \\ &= ACov[x, x]A^T = A\Sigma A^T \end{aligned}$$

$V_x(x+b) = V(x)$
 $Cov(y,y) = E[y y^T] - E[y]E[y]^T$
 $(Ax)^T = x^T A^T \Rightarrow E[Ax]E[Ax]^T = E[x]E[x]^T A^T A$
 $E[Axx^T A^T] = E[x x^T] A^T A$
 $E[x x^T] = E[x]E[x]^T + Cov[x, x]$
 $E[x]E[x]^T = E[x]E[x]^T$
 $E[x x^T] A^T A = E[x]E[x]^T A^T A + Cov[x, x] A^T A$



That is what is used here. Now I can bring out A from the front end and bring out A transpose from the back end, Ok. So we can do that.

So A from the front end, A transpose from the back end and we can take that out common and we get A times expectation of x x transpose minus expectation of x, expectation of x transpose. And A transpose comes out.

So you can now rewrite it. This is nothing but

(Refer Slide Time: 12:45)

Affine transformations


Matrix

- Affine transform : Transformation of variables of the form $y = Ax + b$
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator $E_x[x]$

$$E_y[y] = E_x[Ax + b] = AE_x[x] + b = A\mu + b$$

$$\begin{aligned} V_y[y] &= V_x[Ax + b] \\ &= V_x[Ax] \rightarrow \text{Because } V_x(x+b) = V(x) \\ &= \text{Cov}[Ax, Ax] \\ &= E[(Ax)(Ax)^T] - E[Ax]E[Ax]^T \\ &= E[Axx^T A^T] - AE[x]E[x]^T A^T \\ &= A(E[xx^T] - E[x]E[x]^T)A^T \\ &= ACov[x, x]A^T = A\Sigma A^T \end{aligned}$$

$V_x(x+b) = V(x)$
 $Cov(y,y) = E[y y^T] - E[y]E[y]^T$
 $(Ax)^T = x^T A^T \Rightarrow E[Ax]E[Ax]^T = E[x]E[x]^T A^T A$
 $E[Axx^T A^T] = E[x x^T] A^T A$
 $E[x x^T] = E[x]E[x]^T + Cov[x, x]$
 $E[x]E[x]^T = E[x]E[x]^T$
 $E[x x^T] A^T A = E[x]E[x]^T A^T A + Cov[x, x] A^T A$



covariance of x comma x which we discussed in the previous slide. So A times covariance of x x, x comma x A transpose. Typically the covariance matrix is denoted by

(Refer Slide Time: 13:00)

Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator

$$\mathbb{E}_y[y] = \mathbb{E}_x[Ax + b] = A\mathbb{E}_x[x] + b = A\mu + b$$

$$\begin{aligned} V_y[y] &= V_x[Ax + b] \\ &= V_x[Ax] \rightarrow \text{Because } V_x(x+b) = V(x) \\ &= \text{Cov}[Ax, Ax] \\ &= \mathbb{E}[(Ax)(Ax)^T] - \mathbb{E}[Ax]\mathbb{E}[Ax]^T \\ &= \mathbb{E}[Axx^T A^T] - A\mathbb{E}[x]\mathbb{E}[x]^T A^T \\ &= A(\mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T)A^T \\ &= ACov[x, x]A^T = A\Sigma A^T \end{aligned}$$

$\mu = \mathbb{E}[x]$
 $Cov(x, y) = \mathbb{E}[(y - \mathbb{E}[y])(y - \mathbb{E}[y])^T]$
 $Cov(x, x) = \Sigma = \mathbb{E}[(x - \mathbb{E}[x])(x - \mathbb{E}[x])^T]$

sigma just like, use the expectation of x,

(Refer Slide Time: 13:06)

Affine transformations

- Affine transform : Transformation of variables of the form $y = Ax + b$
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator

$$\mathbb{E}_y[y] = \mathbb{E}_x[Ax + b] = A\mathbb{E}_x[x] + b = A\mu + b$$


$$\begin{aligned} V_y[y] &= V_x[Ax + b] \\ &= V_x[Ax] \rightarrow \text{Because } V_x(x+b) = V(x) \\ &= \text{Cov}[Ax, Ax] \\ &= \mathbb{E}[(Ax)(Ax)^T] - \mathbb{E}[Ax]\mathbb{E}[Ax]^T \\ &= \mathbb{E}[Axx^T A^T] - A\mathbb{E}[x]\mathbb{E}[x]^T A^T \\ &= A(\mathbb{E}[xx^T] - \mathbb{E}[x]\mathbb{E}[x]^T)A^T \\ &= ACov[x, x]A^T = A\Sigma A^T \end{aligned}$$

$\mu = \mathbb{E}[x]$
 $Cov(x, y) = \mathbb{E}[(y - \mathbb{E}[y])(y - \mathbb{E}[y])^T]$
 $Cov(x, x) = \Sigma = \mathbb{E}[(x - \mathbb{E}[x])(x - \mathbb{E}[x])^T]$

capital sigma is the variance of x which is the same as covariance of


(Refer Slide Time: 13:18)

Affine transformations



- Affine transform : Transformation of variables of the form $y = Ax + b$ *Matrix*
- Question: Given the mean and variance matrix of x find them for y
- Then, since expectation is a linear operator
 $E[x]$ $\mu = E[x]$
 $\Sigma = \text{Cov}[x]$

$$E_y[y] = E_x[Ax + b] = AE_x[x] + b = A\mu + b$$

$$\begin{aligned}
 V_y[y] &= V_x[Ax + b] \\
 &= V_x[Ax] \rightarrow \text{Because } V_c(x+b) = V(x) \\
 &= \text{Cov}[Ax, Ax] \\
 &= E[(Ax)(Ax)^T] - E[Ax]E[Ax]^T \quad (\text{Ans})^T = x^T A^T \\
 &= E[Axx^T A^T] - AE[x]E[x]^T A^T \quad E[Ax]^T \\
 &= A(E[xx^T] - E[x]E[x]^T)A^T = E[xx^T] - E[x]E[x]^T \\
 &= ACov[x, x]A^T = A\Sigma A^T \quad \text{Cov}(x, x) = \Sigma = E[xx^T] - E[x]E[x]^T
 \end{aligned}$$


x on x.

So we have looked at a lot of mathematical relations in this video. As I said earlier you might or might not be comfortable with it. Even if you are not comfortable with it, we do not use at least this set of expressions too often. But please get comfortable by watching this a few times in case you found it unfamiliar. We will use it a little bit towards the latter half of the course, thank you.