# Reducción de Dimensionalidad

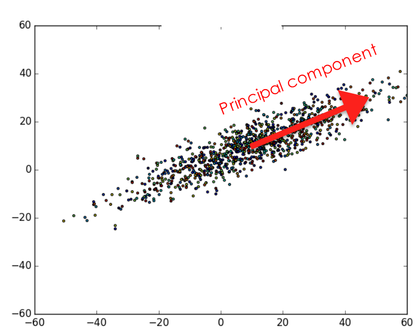
- Se aplican transformaciones de los datos de tal forma de obtener una versión comprimida de ellos.
- La versión comprimida tiene bajos porcentajes de pérdida de información
- Técnica comunmente utilizada: Principal Components Analysis (PCA).

Karim Pichara B.

PUC Chile

# Principal Components Analysis (PCA)

- Método de reducción de dimensionalidad
- Los atributos se transforman en un nuevo conjunto de atributos más reducido, donde cada uno de los atributos del conjunto nuevo es una combinación lineal de los atributos iniciales




PUC Chile

# Recordar algunos conceptos básicos necesarios

- Valor esperado de una variable x (E[x]) en palabras simples es el promedio de esa variable.
- Corresponde al promedio ponderado de ocurrencias de la variable.
- Los ponderadores corresponden a las probabilidades de ocurrencia de cada valor.
- Si todos los valores que puede tomar la variable son equiprobables, el valor esperado corresponde simplemente al promedio de las ocurrencias

# Recordar algunos conceptos básicos necesarios

$$E[x] = \sum_i x_i \, p(x_i)$$ (Caso discreto)

$$E[x] = \int x \, p(x) dx$$ (Caso continuo)

- Si no conocemos p, pero tenemos **N** muestras de p:

$$E[x] \approx \frac{1}{N} \sum_i x_i$$

(Monte Carlo approximation)

$$x_i \sim p$$

muestran aparecen acorde con la distribución p

# Conceptos básicos necesarios

- Valor esperado de una función **f** : (Importante en Machine Learning)

$$E[f(x)] = \int f(x)\, p(x) dx$$

- Si no conocemos p, pero tenemos **N** muestras de p:

$$E[f(x)] \approx \frac{1}{N} \sum_i f(x_i) \quad , \qquad x_i \sim p$$

(Monte Carlo approximation)

Karim Pichara B.

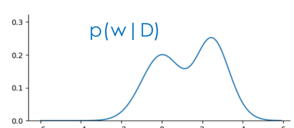
# Conceptos básicos necesarios

- Ejemplo: valor esperado de una función **f**

Modelo predictor de la temperatura en una ciudad

Temp(d) = $\underbrace{\mathbf{w}}_{\text{model}}$ * $\underbrace{\text{Temp (d-1)}}_{\text{data}}$ = $f(w)$

Si w pasa a tener una distribución:




Temp(d) = $\underbrace{\int f(w) * p(w|data)\, dw}_{\text{Expected value of a function}}$

Karim Pichara B.

## Conceptos básicos necesarios

- Ejemplo: valor esperado de una función $\mathbf{f}$

Si la distribución de w no se conoce, pero tenemos N casos:

$$\text{Temp(d)} = \underbrace{\frac{1}{N}\sum_i f(w_i)}_{\text{Monte Carlo Approximation}} \qquad w_i \sim p$$

Los casos vienen de la distribución p



## Recordar algunos conceptos básicos necesarios

- Linealidad:

$$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$$

$$\mathrm{E}[aX] = a\,\mathrm{E}[X]$$

## Recordar algunos conceptos básicos necesarios

- Varianza: Indicador que mide el grado de dispersión de los datos.
  - Corresponde al promedio de las diferencias al cuadrado con la media (el promedio de los datos)

$$\mathrm{Var}(X) = \overset{\text{promedio}}{\mathrm{E}}\left[(X - \overset{\text{promedio}}{\mathrm{E}}[X])^2\right]$$

$$\begin{aligned} &= \mathrm{E}\left[X^2 - 2X\,\mathrm{E}[X] + \mathrm{E}[X]^2\right] \\ &= \mathrm{E}\left[X^2\right] - 2\,\mathrm{E}[X]\,\mathrm{E}[X] + \mathrm{E}[X]^2 \\ &= \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2 \end{aligned}$$

Karim Pichara B. PUC Chile

## Recordar algunos conceptos básicos necesarios

- Covarianza:

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$

$$= \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y]$$

- Matriz de Covarianza: $\mu_i = \mathrm{E}(X_i)$

$$\Sigma = \begin{bmatrix} \mathrm{E}[(X_1-\mu_1)(X_1-\mu_1)] & \mathrm{E}[(X_1-\mu_1)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_1-\mu_1)(X_n-\mu_n)] \\ \mathrm{E}[(X_2-\mu_2)(X_1-\mu_1)] & \mathrm{E}[(X_2-\mu_2)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_2-\mu_2)(X_n-\mu_n)] \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{E}[(X_n-\mu_n)(X_1-\mu_1)] & \mathrm{E}[(X_n-\mu_n)(X_2-\mu_2)] & \cdots & \mathrm{E}[(X_n-\mu_n)(X_n-\mu_n)] \end{bmatrix}$$

## Recordar algunos conceptos básicos necesarios

- Covarianza:

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]$$
$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - E(X))(y_i - E(Y))$$

- Matriz de Covarianza (Notación matricial):

D x 1

D x D $\Sigma = \mathbf{E}\left[(\mathbf{X} - \mathbf{E}[\mathbf{X}])(\mathbf{X} - \mathbf{E}[\mathbf{X}])^{\mathrm{T}}\right]$

$$\Sigma = \mathbf{E}(\mathbf{X}\mathbf{X}^{\mathrm{T}}) - \mathbf{E}(\mathbf{X})\,\mathbf{E}(\mathbf{X})^{\mathrm{T}}$$

Karim Pichara B. PUC Chile

## Recordar algunos conceptos básicos necesarios

- Varianza: Algunas propiedades

$$Var(aX + b) = a^2 Var(X)$$

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X,Y)$$

Karim Pichara B. PUC Chile