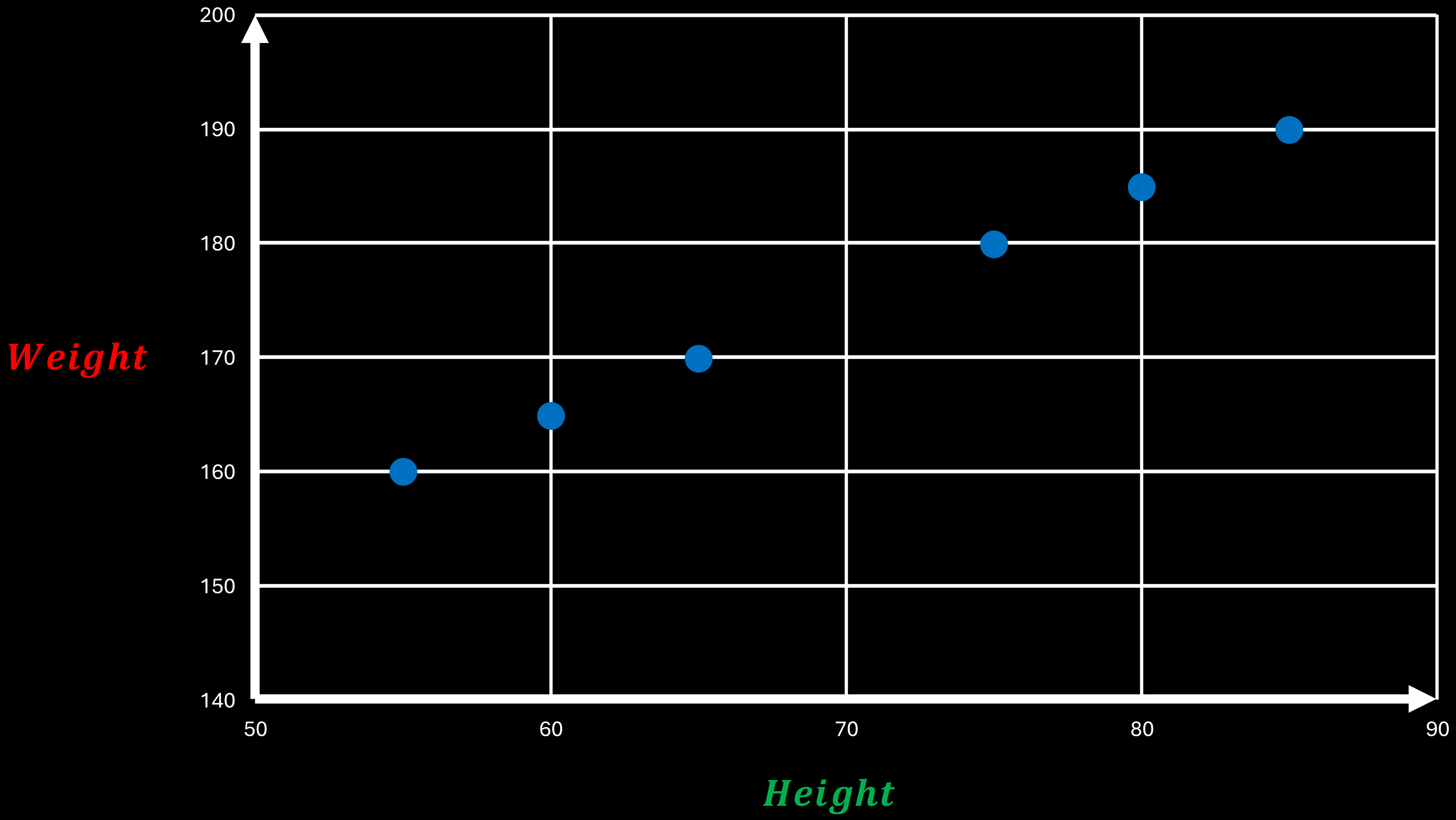
# GMM Example

# Example Dataset

| Person | Height | Weight |
|---|---|---|
| A | 160 | 55 |
| B | 165 | 60 |
| C | 170 | 65 |
| D | 180 | 75 |
| E | 185 | 80 |
| F | 190 | 85 |


200
190
180
170
160
150
140
Weight
50
60
70
80
90
Height

# Initialize cluster parameters

| Cluster | μ (mean) | Σ (covariance) | π (prior) |
|---|---|---|---|
| women | [162, 57] | [[10, 0], [0, 10]] | 0.5 |
| men | [185, 80] | [[10, 0], [0, 10]] | 0.5 |

$$u_{women} = [162, 57]$$

$$\Sigma_{women} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$$

$$\pi_{women} = 0.5$$

$$\mathcal{N}(u_{women}, \Sigma_{women})$$

$$u_{men} = [185, 80]$$

$$\Sigma_{men} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$$

$$\pi_{men} = 0.5$$

$$\mathcal{N}(u_{men}, \Sigma_{men})$$

# Compute the Gaussian density for each sample

$$\boxed{\mathcal{N}(x|\mu,\Sigma) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)}$$

$$\boxed{\mathcal{N}(A|u_{women},\Sigma_{women}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(A-\mu_{women})^T\Sigma_{women}^{-1}(A-\mu_{women})\right)}$$

# Compute the Gaussian density for each sample

$$\boxed{\mathcal{N}(A|u_{women}, \Sigma_{women}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(A-\mu_{women})^T \Sigma_{women}^{-1}(A-\mu_{women})\right)}$$

$$|\Sigma_{women}| = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} = 100$$

Compute determinant of $\Sigma$

$$100^{1/2} = sqrt(100) = 10$$

Convert exponential form to $x^{m/n} = \sqrt[m]{x^n}$

$$2(3.14159)^{2/2} * 10 = 62.83$$

$\pi = 3.14159$

# Compute the Gaussian density for each sample

$$\mathcal{N}(A|u_{women}, \Sigma_{women}) = \frac{1}{62.83} \exp\left(-\frac{1}{2}(A - \mu_{women})^T \Sigma_{women}^{-1} (A - \mu_{women})\right)$$

$$\Sigma_{women}^{-1} = \frac{1}{100}\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$$

Compute inverse of $\Sigma$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{(a \times d) - (b \times c)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

# Compute the Gaussian density for each sample

$$\boxed{\mathcal{N}(A|u_{women}, \Sigma_{women}) = \frac{1}{62.83} \exp\left(-\frac{1}{2}(A-\mu_{women})^T \Sigma_{women}^{-1}(A-\mu_{women})\right)}$$

$$A = \begin{bmatrix} x_h \\ x_w \end{bmatrix} \qquad \mu_{women} = \begin{bmatrix} \mu_h \\ \mu_w \end{bmatrix} \qquad \Sigma_{women}^{-1} = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$$

$$(A-\mu_{women})^T = \left(\begin{bmatrix} x_h \\ x_w \end{bmatrix} - \begin{bmatrix} \mu_h \\ \mu_w \end{bmatrix}\right)^t = \left(\begin{bmatrix} x_h - \mu_h \\ x_w - \mu_w \end{bmatrix}\right)^t = \begin{bmatrix} x_h - \mu_h & x_w - \mu_w \end{bmatrix}$$

$$(A-\mu_{women}) = \begin{bmatrix} x_h \\ x_w \end{bmatrix} - \begin{bmatrix} \mu_h \\ \mu_w \end{bmatrix} = \begin{bmatrix} x_h - \mu_h \\ x_w - \mu_w \end{bmatrix}$$

# Compute the Gaussian density for each sample

$$\boxed{\mathcal{N}(A|u_{women}, \Sigma_{women}) = \frac{1}{62.83}\exp\left(-\frac{1}{2}(A-\mu_{women})^T\Sigma_{women}^{-1}(A-\mu_{women})\right)}$$

$$\Sigma_{women}^{-1}(A-\mu_{women}) = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}\begin{bmatrix} x_h-\mu_h \\ x_w-\mu_w \end{bmatrix} = \begin{bmatrix} 0.1\,(x_h-\mu_h) \\ 0.1\,(x_w-\mu_w) \end{bmatrix}$$

$$(A-\mu_{women})^T\Sigma_{women}^{-1}(A-\mu_{women}) = \begin{bmatrix} x_h-\mu_h & x_w-\mu_w \end{bmatrix}\begin{bmatrix} 0.1\,(x_h-\mu_h) \\ 0.1\,(x_w-\mu_w) \end{bmatrix}$$

$$(A-\mu_{women})^T\Sigma_{women}^{-1}(A-\mu_{women}) = 0.1\,(x_h-\mu_h)^2 + 0.1\,(x_w-\mu_w)^2$$

# Compute the Gaussian density for each sample

$$\boxed{\mathcal{N}(A|u_{women}, \Sigma_{women}) = \frac{1}{62.83} \exp\left(-\frac{1}{2}(A - \mu_{women})^T \Sigma_{women}^{-1}(A - \mu_{women})\right)}$$

$$(A - \mu_{women})^T \Sigma_{women}^{-1}(A - \mu_{women}) = 0.1\,(x_h - \mu_h)^2 + 0.1\,(x_w - \mu_w)^2$$

$$(A - \mu_{women})^T \Sigma_{women}^{-1}(A - \mu_{women}) = \frac{(x_h - \mu_h)^2}{10} + \frac{(x_w - \mu_w)^2}{10}$$

# Compute the Gaussian density for each sample

$$\boxed{\mathcal{N}(A|u_{women}, \Sigma_{women}) = \frac{1}{62.83} \exp\left(-\frac{1}{2}(A-\mu_{women})^T \Sigma_{women}^{-1} (A-\mu_{women})\right)}$$

$$(A-\mu_{women})^T \Sigma_{women}^{-1} (A-\mu_{women}) = \frac{(x_h-\mu_h)^2}{10} + \frac{(x_w-\mu_w)^2}{10}$$

$$(A-\mu_{women})^T \Sigma_{women}^{-1} (A-\mu_{women}) = \frac{(160-162)^2}{10} + \frac{(55-57)^2}{10}$$

$$(A-\mu_{women})^T \Sigma_{women}^{-1} (A-\mu_{women}) = \frac{(-2)^2}{10} + \frac{(-2)^2}{10} = \frac{4}{10} + \frac{4}{10} = \frac{8}{10} = 0.8$$

# Compute the Gaussian density for each sample

$$\boxed{\mathcal{N}(A|u_{women}, \Sigma_{women}) = \frac{1}{62.83} \exp\left(-\frac{1}{2}(A - \mu_{women})^T \Sigma_{women}^{-1} (A - \mu_{women})\right)}$$

$$(A - \mu_{women})^T \Sigma_{women}^{-1} (A - \mu_{women}) \; \frac{(-2)^2}{10} + \frac{(-2)^2}{10} = \frac{4}{10} + \frac{4}{10} = \frac{8}{10} = 0.8$$

$$\mathcal{N}(A|u_{women}, \Sigma_{women}) = \frac{1}{62.83} \exp\left(-\frac{1}{2} \times 0.8\right) = \frac{1}{62.83} \exp(-0.4)$$

$$\boxed{\mathcal{N}(A \mid u_{women}, \Sigma_{women}) = \frac{1}{62.83}(0.6703) = 0.0107}$$

# Compute the Gaussian density for each sample

$$\boxed{\mathcal{N}(A|u_{men}, \Sigma_{men}) = \frac{1}{62.83}\exp\left(-\frac{1}{2}(A-\mu_{men})^T\Sigma_{men}^{-1}(A-\mu_{men})\right)}$$

$$(A-\mu_{men})^T\Sigma_{men}^{-1}(A-\mu_{men}) = \frac{(x_h-\mu_h)^2}{10} + \frac{(x_w-\mu_w)^2}{10}$$

$$(A-\mu_{men})^T\Sigma_{men}^{-1}(A-\mu_{men}) = \frac{(160-185)^2}{10} + \frac{(55-80)^2}{10}$$

$$(A-\mu_{men})^T\Sigma_{men}^{-1}(A-\mu_{men}) = \frac{(-25)^2}{10} + \frac{(-25)^2}{10} = \frac{625}{10} + \frac{625}{10} = \frac{1250}{10} = 125$$

# Compute the Gaussian density for each sample

$$\mathcal{N}(A|u_{men}, \Sigma_{men}) = \frac{1}{62.83} \exp\left(-\frac{1}{2}(A - \mu_{men})^T \Sigma_{men}^{-1} (A - \mu_{men})\right)$$

$$(A - \mu_{men})^T \Sigma_{men}^{-1} (A - \mu_{men}) = \frac{(-25)^2}{10} + \frac{(-25)^2}{10} = \frac{625}{10} + \frac{625}{10} = \frac{1250}{10} = 125$$

$$\mathcal{N}(A|u_{men}, \Sigma_{men}) = \frac{1}{62.83} \exp\left(-\frac{1}{2} \times 125\right) = \frac{1}{62.83} \exp(-62.5)$$

$$\mathcal{N}(A|u_{men}, \Sigma_{men}) = \frac{1}{62.83}(7.2 \times 10^{-28}) = 1.15 \times 10^{-29}$$

# E-step: Compute responsibilities

$$\boxed{r_{ic} = \frac{\pi_c \, \mathcal{N}(x_i \mid \mu_c, \Sigma_c)}{\sum_{j=1}^{K} \pi_J \mathcal{N}\left(x_i \mid \mu_J, \Sigma_J\right)}}$$

$$\boxed{r_{women} = \frac{0.5 \times 0.0107}{0.5 \times 0.0107 + 0.5 \times 0} = 1}$$

$$\boxed{r_{men} = \frac{0.5 \times 0}{0.5 \times 0.0107 + 0.5 \times 0} = 0}$$

# E-step: Compute responsibilities

| Person | Height | Weight | Women | Men |
|---|---|---|---|---|
| A | 160 | 55 | 1 | 0 |
| B | 165 | 60 | | |
| C | 170 | 65 | | |
| D | 180 | 75 | | |
| E | 185 | 80 | | |
| F | 190 | 85 | | |

# E-step: Compute responsibilities

| Person | Height | Weight | Women | Men |
|---|---|---|---|---|
| A | 160 | 55 | 1 | 0 |
| B | 165 | 60 | 1 | 0 |
| C | 170 | 65 | 1 | 0 |
| D | 180 | 75 | 0 | 1 |
| E | 185 | 80 | 0 | 1 |
| F | 190 | 85 | 0 | 1 |

Compute responsibilities $r_{ic}$ for each sample

# M-step: Update parameters

1. $$m_c = \sum_{i=1} r_{ic}$$

Get total number points for each of cluster

3. $$\mu_c = \frac{1}{m_c} \sum_i r_{ic} x^i$$

Update means

2. $$\pi_c = \frac{m_c}{m}$$

Update mixing coefficient

4. $$\Sigma_k = \frac{1}{m_c} \sum_i r_{ic} \left(x^{(i)} - \mu_c\right)^T \left(x^{(i)} - \mu_c\right)$$

Update covariance matrices

# M-step: Update parameters

1

$$m_c = \sum_{i=1} r_{ic}$$

Get total number points for each of cluster

$$m_{women} = 1 + 1 + 1 + 0 + 0 + 0 = 3$$

$$m_{men} = 0 + 0 + 0 + 1 + 1 + 1 = 3$$

# M-step: Update parameters

2

$$\pi_c = \frac{m_c}{m}$$

Update mixing coefficients

$$\boxed{\pi_{women} = \frac{3}{6} = \frac{1}{2} = 0.5}$$

$$\boxed{\pi_{men} = \frac{3}{6} = \frac{1}{2} = 0.5}$$

# M-step: Update parameters

3 $$\mu_c = \frac{1}{m_c} \sum_i r_{ic} x^i$$

Update means

$$\mu_{height} = \frac{160 + 165 + 170}{3} = 165$$

$$\mu_{weight} = \frac{55 + 60 + 65}{3} = 60$$

$$\mu_{women} = \begin{bmatrix} 165 \\ 60 \end{bmatrix}$$

$$\mu_{men} = \frac{180 + 185 + 190}{3} = 185$$

$$\mu_{men} = \frac{75 + 80 + 85}{3} = 80$$

$$\mu_{men} = \begin{bmatrix} 185 \\ 80 \end{bmatrix}$$

# M-step: Update parameters

4

$$\Sigma_c = \frac{1}{m_c} \sum_{n=i} r_{ic} (x_i - \mu_c)^T (x_i - \mu_c)$$

Update covariance matrices

$$\boxed{\mu_{women} = \begin{bmatrix} 165 \\ 60 \end{bmatrix}} \quad A = \begin{bmatrix} 160 \\ 55 \end{bmatrix} \quad B = \begin{bmatrix} 165 \\ 60 \end{bmatrix} \quad C = \begin{bmatrix} 170 \\ 65 \end{bmatrix}$$

$$\boxed{\mu_{men} = \begin{bmatrix} 185 \\ 80 \end{bmatrix}} \quad D = \begin{bmatrix} 180 \\ 75 \end{bmatrix} \quad E = \begin{bmatrix} 185 \\ 80 \end{bmatrix} \quad F = \begin{bmatrix} 190 \\ 85 \end{bmatrix}$$

# M-step: Update parameters

$$\boxed{\mu_{women} = \begin{bmatrix}165\\60\end{bmatrix}} \qquad A = \begin{bmatrix}160\\55\end{bmatrix} \qquad B = \begin{bmatrix}165\\60\end{bmatrix} \qquad C = \begin{bmatrix}170\\65\end{bmatrix}$$

$$(A - \mu_{women})\left(\begin{bmatrix}160\\55\end{bmatrix} - \begin{bmatrix}165\\60\end{bmatrix}\right) = \begin{bmatrix}-5\\-5\end{bmatrix} \qquad (A - \mu_{women})^T = [-5, -5]$$

$$(A - \mu_{women})(A - \mu_{women})^T = \begin{bmatrix}-5\\-5\end{bmatrix}[-5, -5] = \begin{bmatrix}25 & 25\\25 & 25\end{bmatrix}$$

# M-step: Update parameters

$$\boxed{\mu_{women} = \begin{bmatrix} 165 \\ 60 \end{bmatrix}} \qquad A = \begin{bmatrix} 160 \\ 55 \end{bmatrix} \qquad B = \begin{bmatrix} 165 \\ 60 \end{bmatrix} \qquad C = \begin{bmatrix} 170 \\ 65 \end{bmatrix}$$

$$(B - \mu_{women})\left(\begin{bmatrix} 165 \\ 60 \end{bmatrix} - \begin{bmatrix} 165 \\ 60 \end{bmatrix}\right) = [0] \qquad (B - \mu_{women})^T = [0]$$

$$(B - \mu_{women})(B - \mu_{women})^T = [0]$$

# M-step: Update parameters

$$\boxed{\mu_{women} = \begin{bmatrix} 165 \\ 60 \end{bmatrix}} \qquad A = \begin{bmatrix} 160 \\ 55 \end{bmatrix} \qquad B = \begin{bmatrix} 165 \\ 60 \end{bmatrix} \qquad C = \begin{bmatrix} 170 \\ 65 \end{bmatrix}$$

$$(C - \mu_{women}) \left( \begin{bmatrix} 170 \\ 65 \end{bmatrix} - \begin{bmatrix} 165 \\ 60 \end{bmatrix} \right) = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \qquad (C - \mu_{women})^T = [5, 5]$$

$$(C - \mu_{women})(C - \mu_{women})^T = \begin{bmatrix} 5 \\ 5 \end{bmatrix} [5, 5] = \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix}$$

# M-step: Update parameters

$$\Sigma_c = \frac{1}{m_c}\sum_{n=i} r_{ic}\,(x_i - \mu_c)^T(x_i - \mu_c)$$

$$\Sigma_{women} = \frac{1}{3}\begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix} + [0] + \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix}$$

$$\Sigma_{women} = \frac{1}{3}\begin{bmatrix} 50 & 50 \\ 50 & 50 \end{bmatrix}$$

$$\boxed{\Sigma_{women} = \begin{bmatrix} 16.67 & 16.67 \\ 16.67 & 16.67 \end{bmatrix}}$$

# M-step: Update parameters

$$\boxed{\mu_{men} = \begin{bmatrix} 185 \\ 80 \end{bmatrix}} \qquad D = \begin{bmatrix} 180 \\ 75 \end{bmatrix} \qquad E = \begin{bmatrix} 185 \\ 80 \end{bmatrix} \qquad F = \begin{bmatrix} 190 \\ 85 \end{bmatrix}$$

$$(D - \mu_{men}) \left( \begin{bmatrix} 180 \\ 75 \end{bmatrix} - \begin{bmatrix} 185 \\ 80 \end{bmatrix} \right) = \begin{bmatrix} -5 \\ -5 \end{bmatrix} \qquad (D - \mu_{men})^T = [-5, -5]$$

$$(D - \mu_{men})(A - \mu_{men})^T = \begin{bmatrix} -5 \\ -5 \end{bmatrix} [-5, -5] = \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix}$$

# M-step: Update parameters

$$\boxed{\mu_{men} = \begin{bmatrix} 185 \\ 80 \end{bmatrix}} \qquad D = \begin{bmatrix} 180 \\ 75 \end{bmatrix} \qquad E = \begin{bmatrix} 185 \\ 80 \end{bmatrix} \qquad F = \begin{bmatrix} 190 \\ 85 \end{bmatrix}$$

$$(E - \mu_{men}) \left( \begin{bmatrix} 185 \\ 80 \end{bmatrix} - \begin{bmatrix} 185 \\ 80 \end{bmatrix} \right) = [0] \qquad (E - \mu_{men})^T = [0]$$

$$(E - \mu_{men})(E - \mu_{men})^T = [0]$$

# M-step: Update parameters

$$\boxed{\mu_{men} = \begin{bmatrix} 185 \\ 80 \end{bmatrix}} \qquad D = \begin{bmatrix} 180 \\ 75 \end{bmatrix} \qquad E = \begin{bmatrix} 185 \\ 80 \end{bmatrix} \qquad F = \begin{bmatrix} 190 \\ 85 \end{bmatrix}$$

$$(F - \mu_{men}) \left( \begin{bmatrix} 190 \\ 85 \end{bmatrix} - \begin{bmatrix} 185 \\ 80 \end{bmatrix} \right) = \begin{bmatrix} 5 \\ 5 \end{bmatrix} \qquad (F - \mu_{men})^T = [5, 5]$$

$$(F - \mu_{men})(F - \mu_{men})^T = \begin{bmatrix} 5 \\ 5 \end{bmatrix} [5, 5] = \begin{bmatrix} 25 & 25 \\ 25 & 25 \end{bmatrix}$$

# M-step: Update parameters

$$\Sigma_c = \frac{1}{m_c}\sum_{n=i} r_{ic}\,(x_i - \mu_c)^T(x_i - \mu_c)$$

$$\Sigma_{men} = \frac{1}{3}\begin{bmatrix}25 & 25\\25 & 25\end{bmatrix} + [0] + \begin{bmatrix}25 & 25\\25 & 25\end{bmatrix}$$

$$\Sigma_{men} = \frac{1}{3}\begin{bmatrix}50 & 50\\50 & 50\end{bmatrix}$$

$$\boxed{\Sigma_{men} = \begin{bmatrix}16.67 & 16.67\\16.67 & 16.67\end{bmatrix}}$$

# M-step: Update parameters

1

$$m_{women} = 3$$

$$m_{men} = 3$$

Get total number of points for each cluster

2

$$\pi_{women} = 0.5$$

$$\pi_{men} = 0.5$$

Updated mixing coefficients

3

$$\mu_{women} = \begin{bmatrix} 165 \\ 60 \end{bmatrix}$$

$$\mu_{men} = \begin{bmatrix} 185 \\ 80 \end{bmatrix}$$

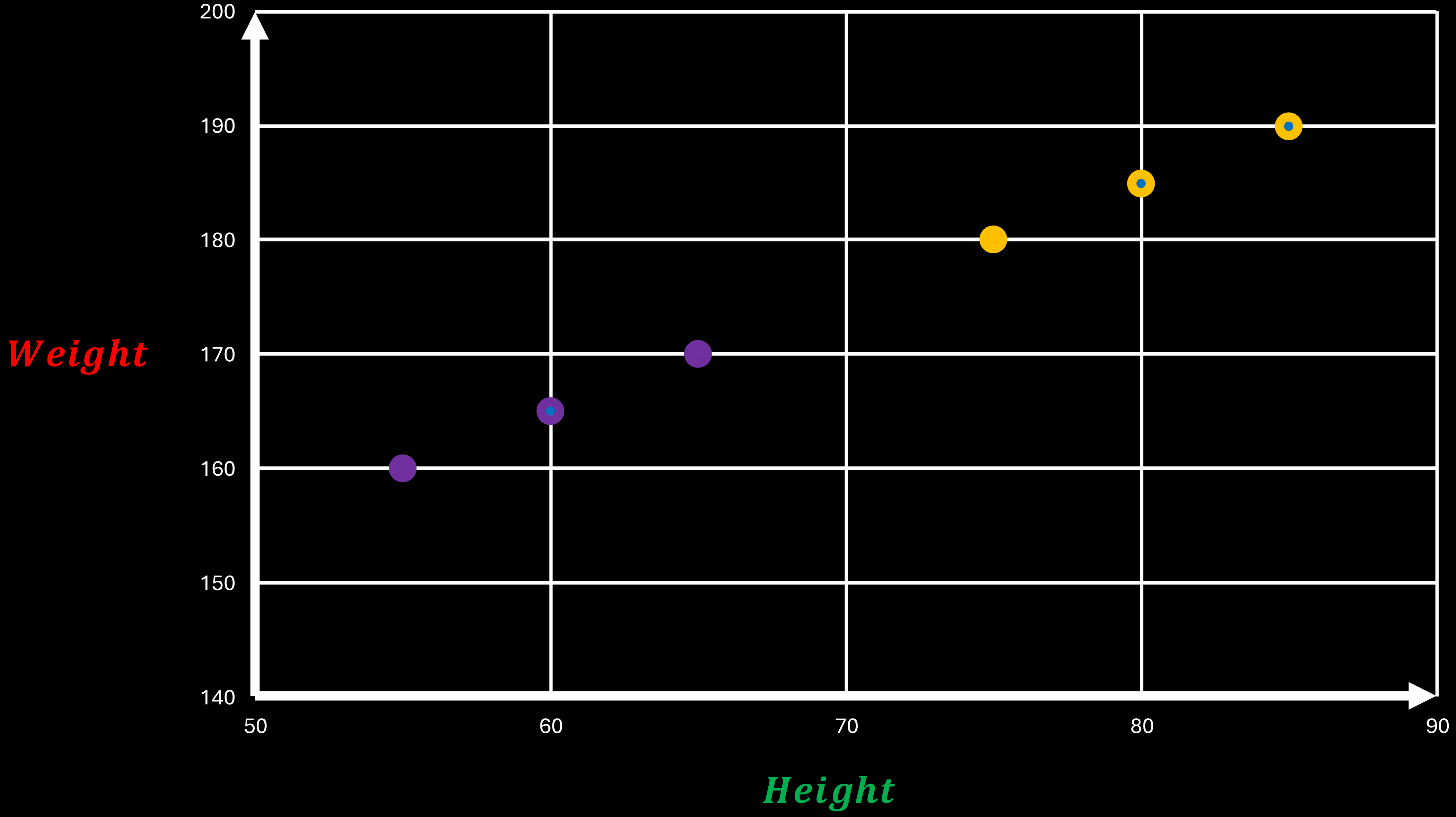
Updated means

4

$$\Sigma_{women} = \begin{bmatrix} 16.67 & 16.67 \\ 16.67 & 16.67 \end{bmatrix}$$

$$\Sigma_{men} = \begin{bmatrix} 16.67 & 16.67 \\ 16.67 & 16.67 \end{bmatrix}$$

Updated covariance matrices


200
190
180
170
160
150
140
50
60
70
80
90
Weight
Height

# New cluster parameters

| Cluster | μ (mean) | Σ (covariance) | π (prior) |
|---|---|---|---|
| women | [165,60] | [[16.67, 16.67], [16.67, 16.67]] | 0.5 |
| men | [185, 80] | [[16.67, 16.67], [16.67, 16.67]] | 0.5 |

$$u_{women} = [162, 57]$$

$$\Sigma_{women} = \begin{bmatrix} 16.67 & 16.67 \\ 16.67 & 16.67 \end{bmatrix}$$

$$\pi_{women} = 0.5$$

$$\mathcal{N}(u_{women}, \Sigma_{women})$$

$$u_{men} = [185, 80]$$

$$\Sigma_{men} = \begin{bmatrix} 16.67 & 16.67 \\ 16.67 & 16.67 \end{bmatrix}$$

$$\pi_{men} = 0.5$$

$$\mathcal{N}(u_{men}, \Sigma_{men})$$