Tutorial 1

- this week

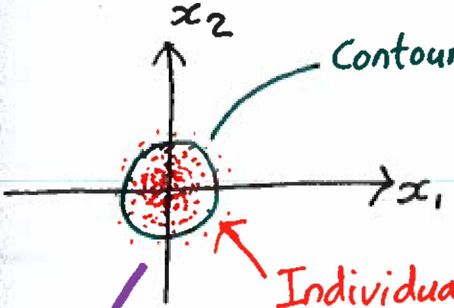
Tutorial 2

- next week

Start assignment 1

Multivariate Gaussians

$$x_d \sim \mathcal{N}(0,1) \Rightarrow p(\underline{x}) = \frac{1}{(2\pi)^{D/2}} e^{\underbrace{-\underline{x}^\top \underline{x}}_{\text{square radius}}/2}$$



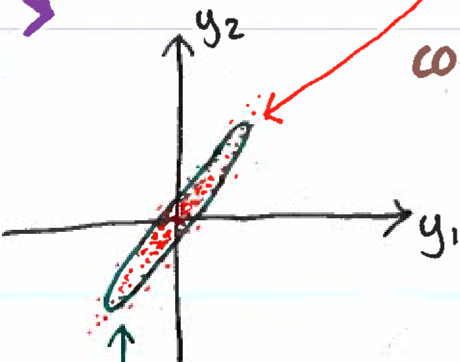
Contours of $p(\underline{x})$

circular/spherical

Individual samples $\underline{x}^{(n)}$

$$\mathrm{cov}[\underline{x}] = \mathbb{I} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

$$\underline{y}^{(n)} = A\,\underline{x}^{(n)}$$ or $\underline{y}^{(n)}$



Elliptical or Ellipsoidal contours.

$$\mathrm{cov}[\underline{y}] = \mathbb{E}[\underline{y}\underline{y}^\top] - \mathbb{E}[\underline{y}]\,\mathbb{E}[\underline{y}]^\top$$

$$= AA^\top = \underbrace{\Sigma}_{\text{Just a common symbol for a covariance.}}$$

PDF of y, $N(y; \underline{0}, \Sigma)$

$y = A\underline{x}$

$\underline{x} = A^{-1}y$

(If A^{-1} exists)

$$p(y) \propto e^{-\frac{1}{2}(A^{-1}y)^T(A^{-1}y)}$$

$$\propto e^{-1/2\, y^T \underbrace{A^{-T}A^{-1}}\, y}$$

$\Sigma = AA^T$

precision $\Sigma^{-1} = A^{-T}A^{-1}$

$$\propto e^{-1/2\, y^T \Sigma^{-1} y}$$

We need $\int p(y)\,dy = 1$

$\underline{x}$'s $\qquad$ y's

volume changes by $|A| = |\Sigma|^{1/2}$

$|\Sigma| = |AA^T| = |A||A^T| = |A|^2$

$$p(y) = \frac{1}{(2\pi)^{D/2}|\Sigma|^{1/2}} e^{-1/2\, y^T \Sigma^{-1} y} = N(y; \underline{0}, \Sigma)$$

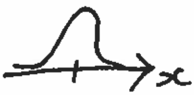
General Gaussian

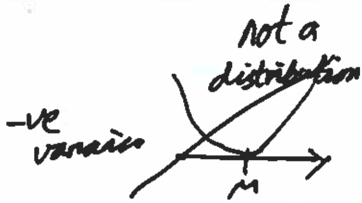
$$\underline{z} = \underbrace{A\underline{x}}_{\underline{y}} + \underline{\mu} \qquad\qquad (\underline{y} = \underline{z} - \underline{\mu})$$

$$p(\underline{z}) = N(\underline{z}; \underline{\mu}, \Sigma)$$

$$= \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} e^{-1/2 (\underline{z}-\underline{\mu})^T \Sigma^{-1} (\underline{z}-\underline{\mu})}$$

Variances are +ve





Covariances usually +ve definite

Iff $\Sigma = AA^T$, Σ is symmetric, positive <u>semi</u>-definite

Positive definite

$$\underline{z}^T \Sigma \underline{z} > 0 \quad \text{for all } \underline{z} \neq 0$$

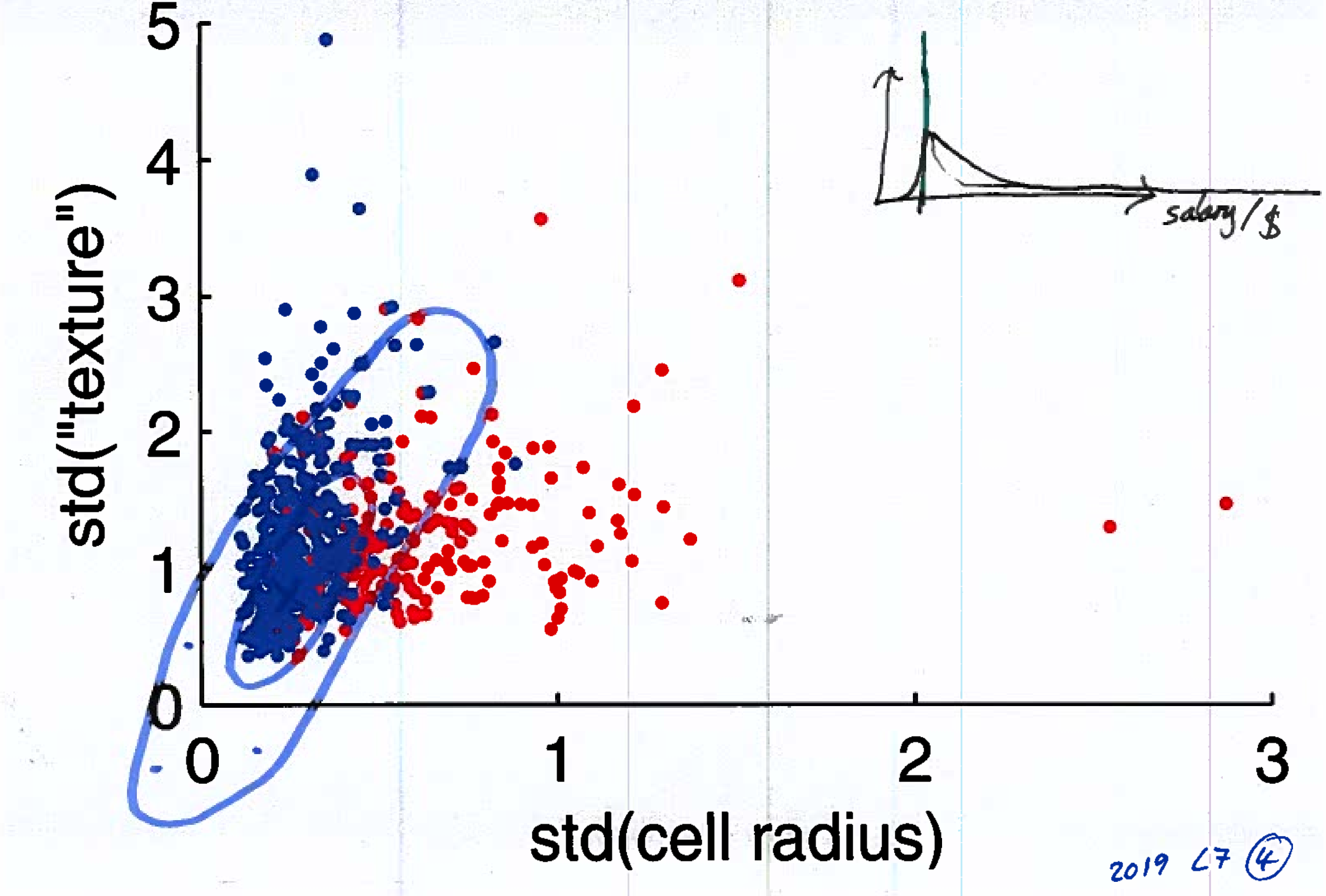
$$\Longleftrightarrow \quad \underline{z}^T \Sigma^{-1} \underline{z} > 0 \quad \text{" " "}$$

[True if A invertible]

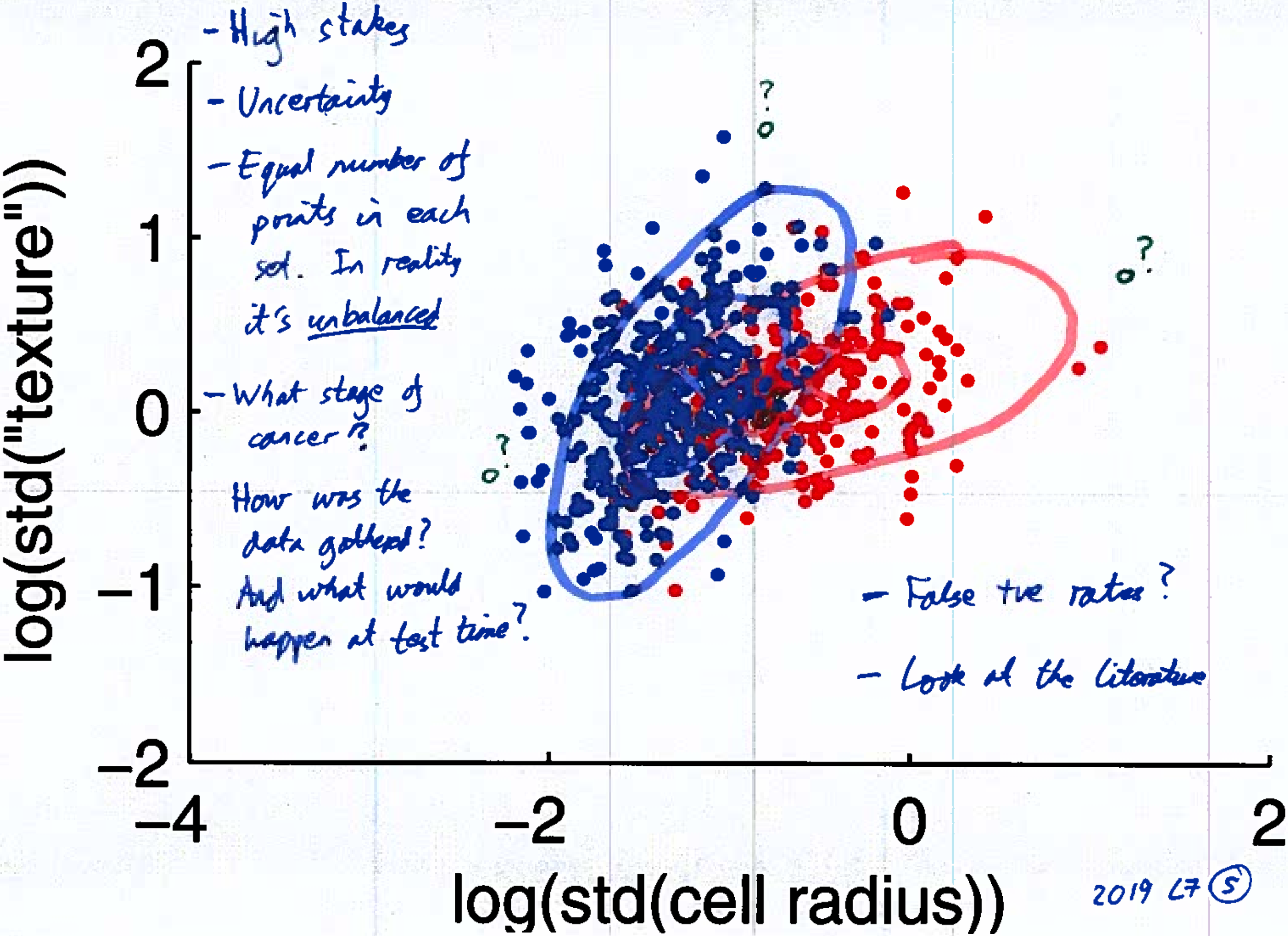
+ve <u>Semi</u>-definite

$$\underline{z}^T \Sigma \underline{z} \geqslant 0 \quad \text{for all } \underline{z}$$

Example $\Sigma = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$



5
4
3
2
1
0
std("texture")
0
1
2
3
std(cell radius)
salary/$
2019 L7 ④


- High stakes
- Uncertainty
- Equal number of points in each set. In reality it's unbalanced
- What stage of cancer?
How was the data gathered? And what would happen at test time?
- False +ve rates?
- Look at the literature
log(std("texture"))
2
1
0
−1
−2
−4
−2
0
2
log(std(cell radius))
2019 L7 5

Bayes Classifier

At training time create model:

Features $p(\underline{x} \mid y=k)$ eg. $\mathcal{N}(\underline{x}; \underline{\mu}^{(k)}, \Sigma^{(k)})$

(or discrete distribution

or Naive Bayes Classifier:
features $\{x_d\}$ are independent
given class)

Labels $p(y=k) = \pi_k$

$\left(\sum_k \pi_k = 1\right)$

Fit the model $\pi_k \approx \dfrac{\#\text{ points in class } k}{N}$

(Or domain knowledge)

$\underline{\mu}^{(k)}, \Sigma^{(k)}$ set mean and covariance
of points in class k.

At test time use Bayes' rule

$$p(y|\underline{x}) \propto p(y)\,p(\underline{x}|y)$$