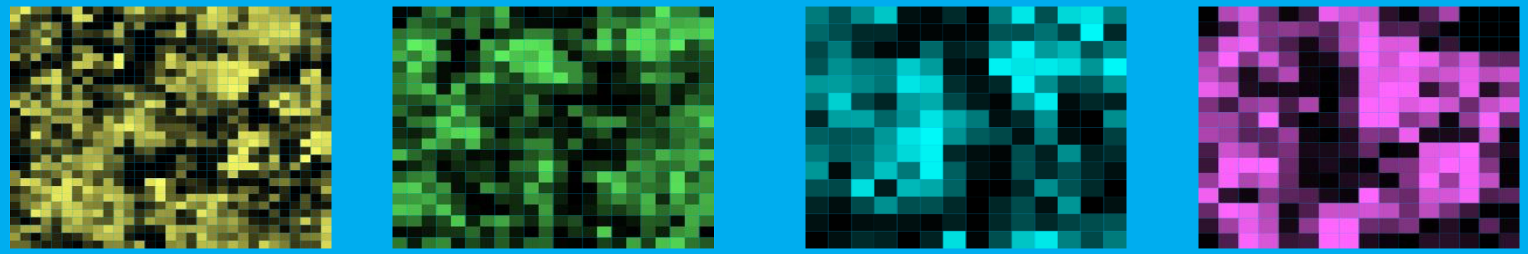
The Characteristic Direction: A Geometrical Approach to Differential Expression – Part Two



Network Analysis in Systems Biology

Neil Clark, PhD

Instructor, Ma’ayan Lab

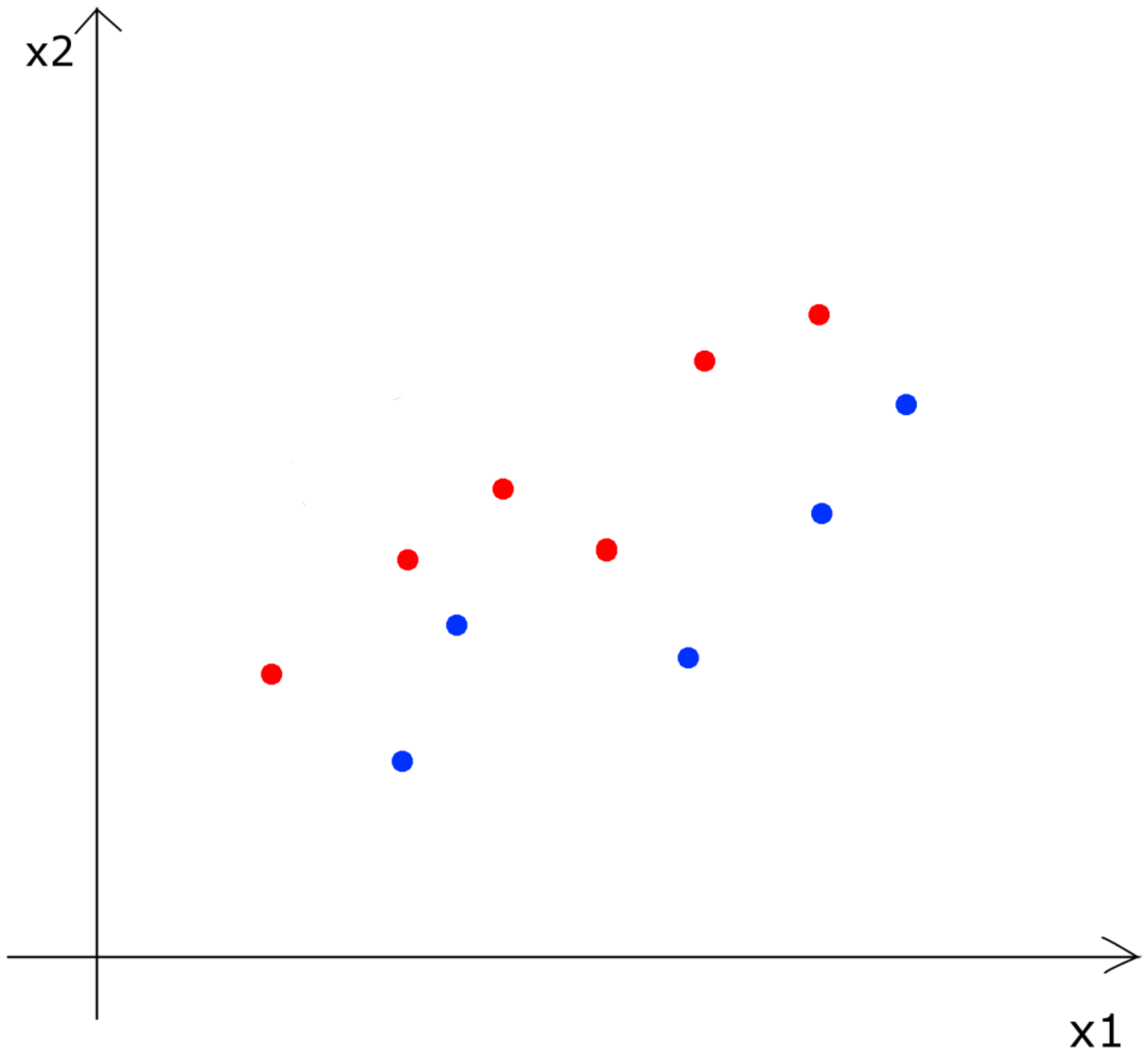
Department of Pharmacology and Systems Therapeutics

Icahn School of Medicine at Mount Sinai, New York, NY 10029

Linear Discriminant Analysis

- Bayes rule for the classification probability:

$$\Pr(G = k|X = x) = \frac{\overbrace{f_k(x)}^{\text{Class conditional density of X}}\overbrace{\pi_k}^{\text{Prior probability of class k}}}{\sum_{l=1}^{K} f_l(x)\pi_l}$$



Linear Discriminant Analysis

- Bayes rule for the classification probability:

Class conditional density of X

Prior probability of class k

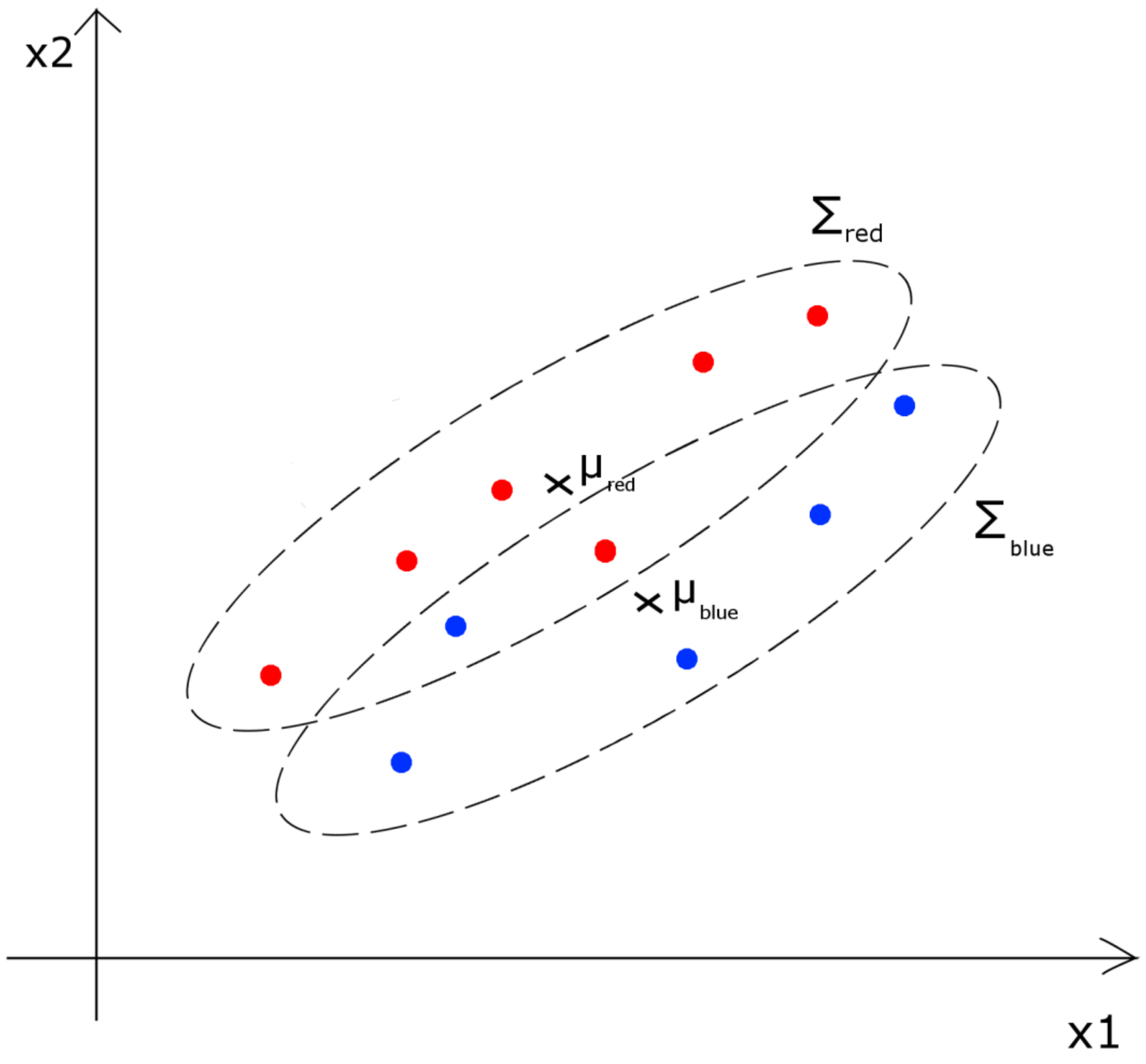
$$\Pr(G = k|X = x) = \frac{f_k(x)\pi_k}{\sum_{l=1}^{K} f_l(x)\pi_l}$$

- Model the class conditional density:

Class mean

Class covariance matrix

$$f_k(x) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)}$$



Linear Discriminant Analysis

- Bayes rule for the classification probability:

Class conditional density of X

Prior probability of class k

$$\Pr(G = k|X = x) = \frac{f_k(x)\pi_k}{\sum_{l=1}^{K} f_l(x)\pi_l}$$

- Model the class conditional density:

Class mean

Class covariance matrix

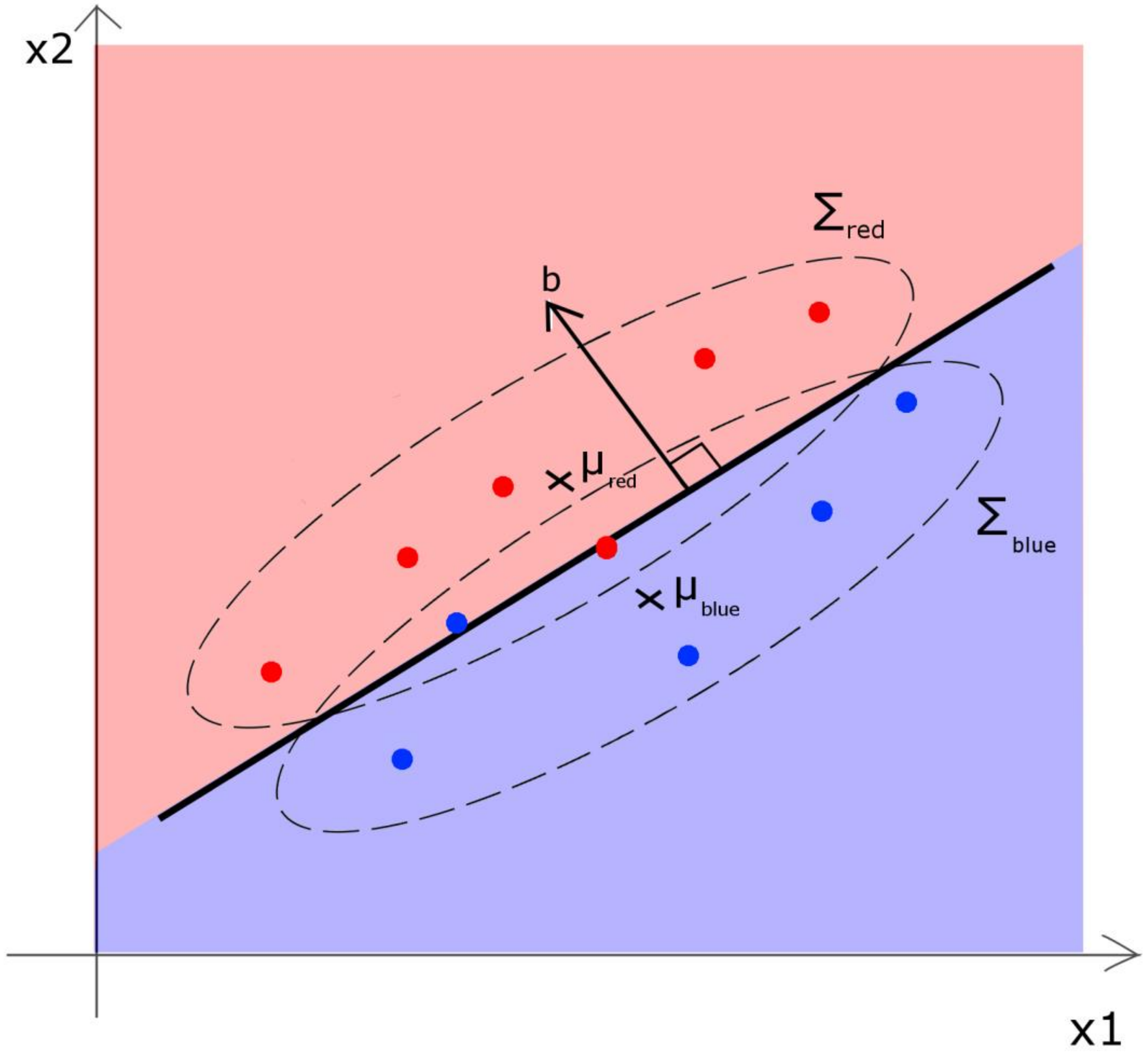
$$f_k(x) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)}$$

- Estimate classification probability:

$$log\frac{\Pr(G = k|X = x)}{\Pr(G = l|X = x)} = log\frac{\pi_k}{\pi_l} - \frac{1}{2}(\mu_k - \mu_l)^T\Sigma^{-1}(\mu_k - \mu_l) + x^T\Sigma^{-1}(\mu_k - \mu_l)$$

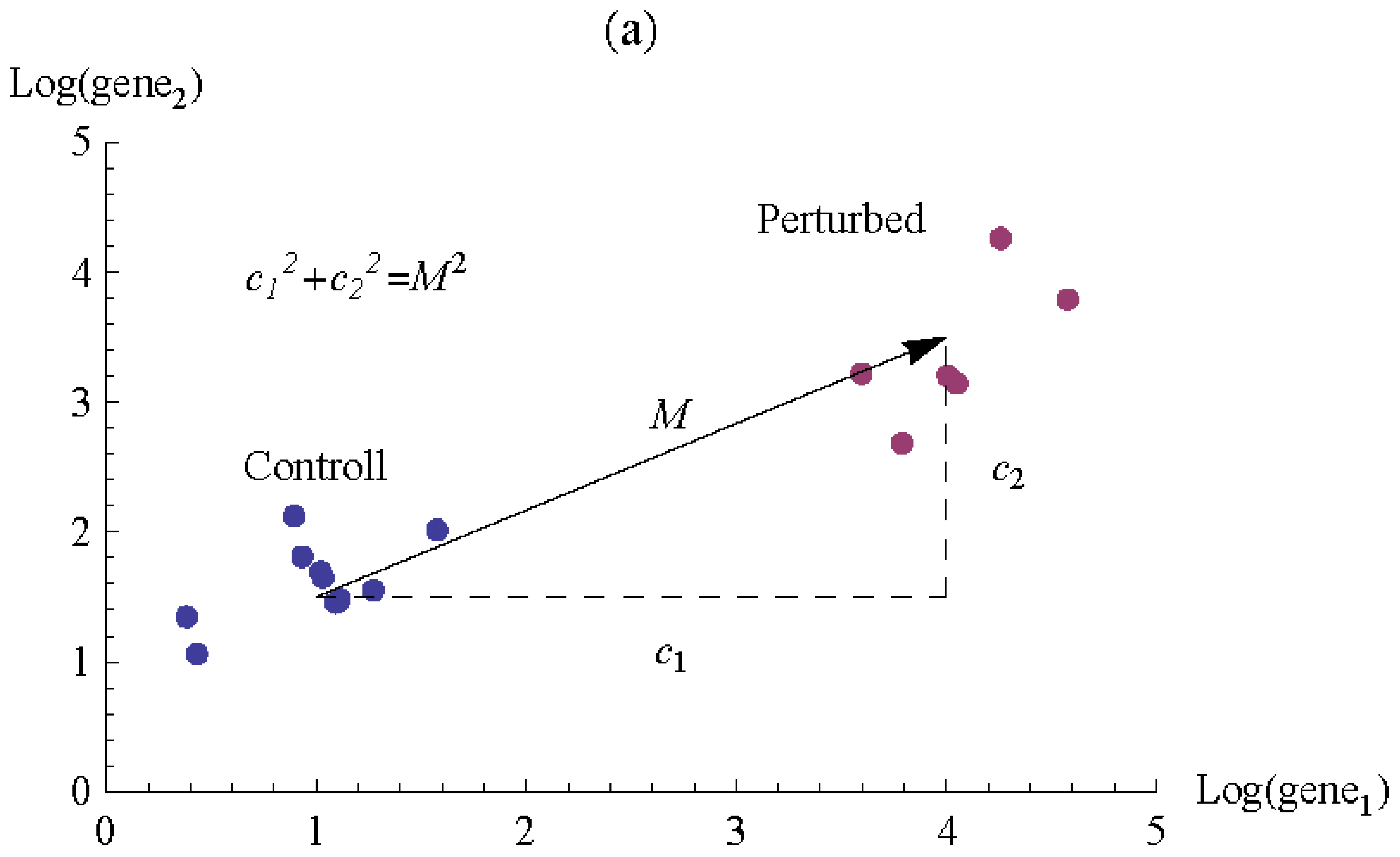
- The orientation of the separating hyper plane

  is given by: $b = \Sigma^{-1}(\mu_k - \mu_l)$



Interpreting the Characteristic Direction

- The square of the component corresponding to each gene is interpreted as a quantification of its significance in the total differential *M*



$$c_1^2 + c_1^2 + c_3^2 + \ldots c_n^2 = M^2$$

Use Shrinkage to Address the Curse of Dimensionality

- Our estimate of the covariance matrix is likely to be fraught with error because of the curse of dimensionality
- Attempt to smooth away the error while retaining the signal with shrinkage
- Shrink the covariance matrix to the scalar variance
  - $\hat{\Sigma}(\gamma) = \gamma\hat{\Sigma} + (1-\gamma)\sigma^2 I_p,$ with $\gamma \epsilon [0,1]$