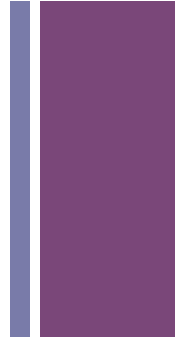
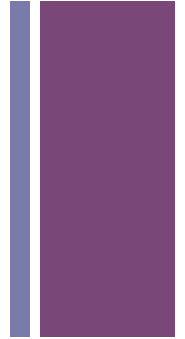


+ Canonical correlation analysis



- A method of correlating linear relationships between two multidimensional variables.
 - $\mathbf{x} = [x_1, x_2, \dots, x_{q_1}]$, $\mathbf{y} = [y_1, y_2, \dots, y_{q_2}]$
 - Transform $\mathbf{x} = \mathbf{w}_x \mathbf{x} = [w_x x_1, w_x x_2, \dots, w_x x_{q_1}]$
 - Transform $\mathbf{y} = \mathbf{w}_y \mathbf{y} = [w_y y_1, w_y y_2, \dots, w_y y_{q_2}]$
 - Choose \mathbf{w}_x and \mathbf{w}_y that maximize the correlation between the transformed vectors

+ Canonical correlation analysis



- Extension of multiple regression – more than one y
 - $x = [x_1, x_2, \dots, x_{q_1}]$, $y = [y_1, y_2, \dots, y_{q_2}]$
 - R_{11} = cor matrix of variables in x
 - R_{22} = cor matrix of variables in y
 - R_{12} = cor between x and y (q_1 by q_2 matrix)
 - $E_1 = R_{11}^{-1} R_{12} R_{22}^{-1} R_{21}$
- CC combines the DVs to find pairs of new variables (called canonical variables, one for each data table) which have the highest correlation.
- CV's, even when highly correlated, do not necessarily explain a large portion of the variance of the original tables.
 - This make the interpretation of the CV sometimes difficult, but CC is nonetheless an important theoretical tool