Análise de correspondência

Prof. Caio Azevedo

13 de março de 2012

- Foram entrevistados 1320 consumidores de aparelhos de som.
- Variáveis medidas: marca adquirida e principal motivo da compra.
- Os dados obtidos encontram-se na Tabela abaixo:

	Atributo						
Marca	**Qual. de som**	**Tec.**	**Pot. do som**	**Rec. técnicos**	**Preços**	**Conf. na marca**	**Total**
Sony	135	140	95	55	40	60	525
Aiwa	50	115	40	60	5	15	285
Gradiente	90	55	20	35	40	10	250
Philips	60	25	35	10	5	30	165
Sharp	30	20	5	10	10	20	95
Total	365	355	195	170	100	135	1320

- Tabelas de contingência.

	Variável 2						
Variável 1	**Cat. 1**	**Cat. 2**	**Cat. 3**	...	**Cat. J**	...	**Total**
Categoria 1	X_{11}	X_{12}	X_{13}	...	X_{1J}	...	$X_{1.}$
Categoria 2	X_{21}	X_{22}	X_{23}	...	X_{2J}	...	$X_{2.}$
Categoria 3	X_{31}	X_{32}	X_{33}	...	X_{3J}	...	$X_{3.}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Categoria i	X_{i1}	X_{i2}	X_{i3}	$\vdots$	X_{iJ}	$\vdots$	$X_{i.}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Categoria I	X_{I1}	X_{I2}	X_{I3}	$\vdots$	X_{IJ}	$\vdots$	$X_{I.}$
Total	$X_{.1}$	$X_{.2}$	$X_{.3}$	...	$X_{.J}$	...	$X_{..}$

- Seja $X_{ij} = 1$ se o indivíduo foi classificado (pertence) na categoria i da variável 1 e na categoria j da variável 2.
- Seja $p_{ij} = P(X_{ij} = 1)$.
- Sob algumas suposições e para $n = X_{..}$ fixado, temos que o número de indivíduos observados nas $I \times J$ categorias seguem uma distribuição multinomial n, p_{ij}.
- Hipótese de interesse: H_0 : as variáveis são estatisticamente independentes vs H_1 : as variáveis não são estatísticamente independentes.
- $H_0 : p_{ij} = p_{i.}p_{.j}, \forall i, j$ vs $H_1 : p_{ij} \neq p_{i.}p_{.j}$, para pelo menos um par (i, j).
- $p_{i.}$ e $p_{.j}$ são as probabilidades marginais de cada indivíduo pertencer, respectivamente à categoria i da variável 1 e à categoria j da variável 2.

- $p_{i.} = \sum_{j=1}^{J} p_{ij}$ e $p_{.j} = \sum_{i=1}^{I} p_{ij}$.
- Sob a independência, temos que $Q = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{(X_{ij} - E_{ij})^2}{E_{ij}}$, em que $E_{ij} = X_{i.}X_{.j}/X_{..}$, apresentará um valor baixo.
- E_{ij} é a freqüência (valor esperado) da casela (i, j) sob independência.
- Sob a validade de suposição de independência, $Q \approx \chi^2_{(I-1)(J-1)}$

Análise de correspondência

Prof. Caio Azevedo

Estimação por máxima verossimilhança)

Análise de correspondência: exemplo)

Teste de qui-quadrado)

Teste de qui-quadrado)

Objetivos)

Análise de correspondência)

Definições)

Definições)

Definições)

Análise de correspondência: exemplo)

Matrizes de

- Verificar se existe associação entre marca e atributo em termos de aquisição por parte dos consumidores.
- Estatística de qui-quadrado: $Q = 179,62$, p-valor $< 0,001$.
- Investigar mais minunciosamente a relação de dependência entre as variáveis que definem a tabela de contingência.
- Por exemplo: que marca de aparelho de som é mais adquirida em função do preço?

- Análise de correspondência: visa medir o grau de associação de variáveis categorizadas dispostas em tabelas de contingência.
- A disposição dos resultados é feita de modo gráfico.
- Uma forma de medir associação é através da estatística de qui-quadrado.
- Alternativa: modelos lineares, log-lineares e não lineares para dados categorizados.

- Note que $Q = (\mathbf{X}^* - \mathbf{E}^*)' \mathbf{D}_{\mathbf{E}^*}^{-1} (\mathbf{X}^* - \mathbf{E}^*)$.
- $\mathbf{X}^* = vec(\mathbf{X}')$ e $\mathbf{E}^* = (E_{11}, ...E_{IJ})'$.
- Defina $\mathbf{P} = \mathbf{X}/n = \mathbf{X}/X_{..}$.
- $\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & \ldots & P_{1J} \\ P_{21} & P_{22} & \ldots & P_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ P_{I1} & P_{I2} & \ldots & P_{IJ} \end{bmatrix}$
- $\mathbf{P}_E = \begin{bmatrix} E_{11} & E_{12} & \ldots & E_{1J} \\ E_{21} & E_{22} & \ldots & E_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ E_{I1} & E_{I2} & \ldots & E_{IJ} \end{bmatrix}$
- $X_{i.} = \sum_{j=1}^{J} X_{ij}$, $X_{.j} = \sum_{i=1}^{I} X_{ij}$ e $X_{..} = \sum_{i=1}^{I} \sum_{j=1}^{J} X_{ij}$.

- Pode-se demonstrar que $Q = n\left(\mathbf{P}-\mathbf{P}_E\right)'\mathbf{D}_{\mathbf{P}_E}^{-1}\left(\mathbf{P}-\mathbf{P}_E\right)$ (exercício).
- A análise de correspondência explora a forma acima.
- Mais especificamente, perfil das linhas e perfil das colunas.
- Perfil das linhas: $p_{j/i} = \frac{p_{ij}}{p_i}$.
- Perfil das linhas: $p_{i/j} = \frac{p_{ij}}{p_j}$.
- $\mathbf{P}_r = \mathbf{P}\mathbf{1}_J$.
- $\mathbf{P}_c = \mathbf{P}'\mathbf{1}_I$.

Análise de correspondência

Prof. Caio Azevedo

Estimação por máxima verossimilhança)

Análise de correspondência: exemplo)

Teste de qui-quadrado)

Teste de qui-quadrado)

Objetivos)

Análise de correspondência)

Definições)

Definições)

Definições)

Análise de correspondência: exemplo)

Matrizes de ...

- $$\mathbf{P}_{r(I\times 1)} = \begin{bmatrix} \sum_{j=1}^{J} P_{1j} \\ \sum_{j=1}^{J} P_{2j} \\ \vdots \\ \sum_{j=1}^{J} P_{Ij} \end{bmatrix} = \begin{bmatrix} P_{1.} \\ P_{2.} \\ \vdots \\ P_{I.} \end{bmatrix} = \begin{bmatrix} \frac{X_{1.}}{X_{..}} \\ \frac{X_{2.}}{X_{..}} \\ \vdots \\ \frac{X_{I.}}{X_{..}} \end{bmatrix}.$$
- $$\mathbf{P}_{c(J\times 1)} = \begin{bmatrix} \sum_{i=1}^{I} P_{i1} \\ \sum_{i=1}^{I} P_{i2} \\ \vdots \\ \sum_{i=1}^{I} P_{iJ} \end{bmatrix} = \begin{bmatrix} P_{.1} \\ P_{.2} \\ \vdots \\ P_{.J} \end{bmatrix} = \begin{bmatrix} \frac{X_{.1}}{X_{..}} \\ \frac{X_{.2}}{X_{..}} \\ \vdots \\ \frac{X_{.I}}{X_{..}} \end{bmatrix}.$$

- Proporções estimadas

	Atributo						
Marca	**Qual. de som**	**Tec.**	**Pot. do som**	**Rec. técnicos**	**Preços**	**Conf. na marca**	**Total**
Sony	0.102	0.106	0.072	0.042	0.030	0.045	0.398
Aiwa	0.038	0.087	0.030	0.045	0.004	0.011	0.216
Gradiente	0.068	0.042	0.015	0.027	0.030	0.008	0.189
Philips	0.045	0.019	0.027	0.008	0.004	0.023	0.125
Sharp	0.023	0.015	0.004	0.008	0.008	0.015	0.072
Total	0.277	0.269	0.148	0.129	0.076	0.102	1.000

- Perfis das linhas: $\mathbf{R} = \mathbf{D}_r^{-1}\mathbf{P} =$

$$\begin{bmatrix} \frac{P_{11}}{P_{1.}} & \frac{P_{12}}{P_{1.}} & \cdots & \frac{P_{1J}}{P_{1.}} \\ \frac{P_{21}}{P_{2.}} & \frac{P_{22}}{P_{2.}} & \cdots & \frac{P_{2J}}{P_{2.}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{P_{I1}}{P_{I.}} & \frac{P_{I2}}{P_{I.}} & \cdots & \frac{P_{IJ}}{P_{I.}} \end{bmatrix} = \begin{bmatrix} \frac{X_{11}}{X_{1.}} & \frac{X_{12}}{X_{1.}} & \cdots & \frac{X_{1J}}{X_{1.}} \\ \frac{X_{21}}{X_{2.}} & \frac{X_{22}}{X_{2.}} & \cdots & \frac{X_{2J}}{X_{2.}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{X_{I1}}{X_{I.}} & \frac{X_{I2}}{P_{I.}} & \cdots & \frac{X_{IJ}}{X_{I.}} \end{bmatrix}$$

- Perfis das colunas: $\mathbf{C} = \mathbf{D}_c^{-1}\mathbf{P}' =$

$$\begin{bmatrix} \frac{P_{11}}{P_{.1}} & \frac{P_{21}}{P_{.1}} & \cdots & \frac{P_{I1}}{P_{.1}} \\ \frac{P_{12}}{P_{.2}} & \frac{P_{22}}{P_{.2}} & \cdots & \frac{P_{I2}}{P_{.2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{P_{1J}}{P_{.J}} & \frac{P_{21}}{P_{.J}} & \cdots & \frac{P_{IJ}}{P_{.J}} \end{bmatrix} = \begin{bmatrix} \frac{X_{11}}{X_{.1}} & \frac{X_{21}}{X_{.1}} & \cdots & \frac{X_{I1}}{X_{.1}} \\ \frac{X_{12}}{X_{.2}} & \frac{X_{22}}{X_{.2}} & \cdots & \frac{X_{I2}}{X_{.2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{X_{1J}}{X_{.J}} & \frac{X_{21}}{X_{.J}} & \cdots & \frac{X_{IJ}}{X_{.J}} \end{bmatrix}$$

- Perfis das linhas

	Atributo						
Marca	**Qual. de som**	**Tec.**	**Pot. do som**	**Rec. técnicos**	**Preços**	**Conf. na marca**	**Total**
Sony	25.71	26.67	18.10	10.48	7.62	11.43	100
Aiwa	17.54	40.35	14.04	21.05	1.75	5.26	100
Gradiente	36.00	22.00	8.00	14.00	16.00	4.00	100
Philips	36.36	15.15	21.21	6.06	3.03	18.18	100
Sharp	31.58	21.05	5.26	10.53	10.53	21.05	100
Total	27.65	26.89	14.77	12.88	7.58	10.23	100

- Perfis das colunas

	Atributo						
Marca	**Qual. de som**	**Tec.**	**Pot. do som**	**Rec. técnicos**	**Preços**	**Conf. na marca**	**Total**
Sony	36.99	39.44	48.72	32.35	40.00	44.44	39.77
Aiwa	13.70	32.39	20.51	35.29	5.00	11.11	21.59
Gradiente	24.66	15.49	10.26	20.59	40.00	7.41	18.94
Philips	16.44	7.04	17.95	5.88	5.00	22.22	12.50
Sharp	8.22	5.63	2.56	5.88	10.00	14.81	7.20
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

- Objetivo: obter uma forma simplificada e interpretável para as seguintes matrizes:

$$\mathbf{W}_{(I \times J)} = \mathbf{D}_r^{1/2} \left(\mathbf{R} - \mathbf{1}\mathbf{P}'_c\right) \mathbf{D}_c^{-1/2}$$

$$\mathbf{Z}_{(J \times I)} = \mathbf{D}_c^{1/2} \left(\mathbf{C} - \mathbf{1}\mathbf{P}'_r\right) \mathbf{D}_r^{-1/2}$$

- Simplificada: duas dimensões (gráfico de dispersão)
- Interpretável: pontos próximos (categorias) tem um maior grau de dependência.

- Note que:

$$\mathbf{R} - \mathbf{1}\mathbf{P}_c' = \begin{bmatrix} \frac{X_{11}}{X_{1.}} - \frac{X_{.1}}{X_{..}} & \frac{X_{12}}{X_{1.}} - \frac{X_{.2}}{X_{..}} & \cdots & \frac{X_{1J}}{X_{1.}} - \frac{X_{.J}}{X_{..}} \\ \frac{X_{21}}{X_{2.}} - \frac{X_{.1}}{X_{..}} & \frac{X_{22}}{X_{2.}} - \frac{X_{.2}}{X_{..}} & \cdots & \frac{X_{2J}}{X_{2.}} - \frac{X_{.J}}{X_{..}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{X_{I1}}{X_{I.}} - \frac{X_{.1}}{X_{..}} & \frac{X_{I2}}{X_{I.}} - \frac{X_{.2}}{X_{..}} & \cdots & \frac{X_{IJ}}{X_{I.}} - \frac{X_{.J}}{X_{..}} \end{bmatrix}$$

$$\mathbf{D}_r^{1/2}\left(\mathbf{R} - \mathbf{1}\mathbf{P}_c'\right) =$$

$$\begin{bmatrix} \sqrt{\frac{X_{1.}}{X_{..}}}\left(\frac{X_{11}}{X_{1.}} - \frac{X_{.1}}{X_{..}}\right) & \sqrt{\frac{X_{1.}}{X_{..}}}\left(\frac{X_{12}}{X_{1.}} - \frac{X_{.2}}{X_{..}}\right) & \cdots & \sqrt{\frac{X_{1.}}{X_{..}}}\left(\frac{X_{1J}}{X_{1.}} - \frac{X_{.J}}{X_{..}}\right) \\ \sqrt{\frac{X_{2.}}{X_{..}}}\left(\frac{X_{21}}{X_{2.}} - \frac{X_{.1}}{X_{..}}\right) & \sqrt{\frac{X_{2.}}{X_{..}}}\left(\frac{X_{22}}{X_{2.}} - \frac{X_{.2}}{X_{..}}\right) & \cdots & \sqrt{\frac{X_{2.}}{X_{..}}}\left(\frac{X_{2J}}{X_{2.}} - \frac{X_{.J}}{X_{..}}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\frac{X_{I.}}{X_{..}}}\left(\frac{X_{I1}}{X_{I.}} - \frac{X_{.1}}{X_{..}}\right) & \sqrt{\frac{X_{I.}}{X_{..}}}\left(\frac{X_{I2}}{X_{I.}} - \frac{X_{.2}}{X_{..}}\right) & \cdots & \sqrt{\frac{X_{I.}}{X_{..}}}\left(\frac{X_{IJ}}{X_{I.}} - \frac{X_{.J}}{X_{..}}\right) \end{bmatrix}$$

- Por outro lado:

$$\mathbf{W} =$$

$$\begin{bmatrix} \sqrt{\frac{X_{1.}}{X_{.1}}}\left(\frac{X_{11}}{X_{1.}} - \frac{X_{.1}}{X_{..}}\right) & \sqrt{\frac{X_{1.}}{X_{.2}}}\left(\frac{X_{12}}{X_{1.}} - \frac{X_{.2}}{X_{..}}\right) & \cdots & \sqrt{\frac{X_{1.}}{X_{.J}}}\left(\frac{X_{1J}}{X_{1.}} - \frac{X_{.J}}{X_{..}}\right) \\ \sqrt{\frac{X_{2.}}{X_{.1}}}\left(\frac{X_{21}}{X_{2.}} - \frac{X_{.1}}{X_{..}}\right) & \sqrt{\frac{X_{2.}}{X_{.2}}}\left(\frac{X_{22}}{X_{2.}} - \frac{X_{.2}}{X_{..}}\right) & \cdots & \sqrt{\frac{X_{2.}}{X_{.J}}}\left(\frac{X_{2J}}{X_{2.}} - \frac{X_{.J}}{X_{..}}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\frac{X_{I.}}{X_{.1}}}\left(\frac{X_{I1}}{X_{I.}} - \frac{X_{.1}}{X_{..}}\right) & \sqrt{\frac{X_{I.}}{X_{.2}}}\left(\frac{X_{I2}}{X_{I.}} - \frac{X_{.2}}{X_{..}}\right) & \cdots & \sqrt{\frac{X_{I.}}{X_{.J}}}\left(\frac{X_{IJ}}{X_{I.}} - \frac{X_{.J}}{X_{..}}\right) \end{bmatrix}$$

$$\begin{bmatrix} \left[X_{11} - \left(\frac{X_{1.}X_{.1}}{X_{..}}\right)\right]\frac{1}{(X_{1.}X_{.1})^{1/2}} & \left[X_{12} - \left(\frac{X_{1.}X_{.2}}{X_{..}}\right)\right]\frac{1}{(X_{1.}X_{.2})^{1/2}} & \cdots & \left[X_{1J} - \left(\frac{X_{1.}X_{.J}}{X_{..}}\right)\right]\frac{1}{(X_{1.}X_{.J})^{1/2}} \\ \left[X_{21} - \left(\frac{X_{2.}X_{.1}}{X_{..}}\right)\right]\frac{1}{(X_{2.}X_{.1})^{1/2}} & \left[X_{22} - \left(\frac{X_{2.}X_{.2}}{X_{..}}\right)\right]\frac{1}{(X_{2.}X_{.2})^{1/2}} & \cdots & \left[X_{2J} - \left(\frac{X_{2.}X_{.J}}{X_{..}}\right)\right]\frac{1}{(X_{2.}X_{.J})^{1/2}} \\ \vdots & \vdots & \ddots & \vdots \\ \left[X_{I1} - \left(\frac{X_{I.}X_{.1}}{X_{..}}\right)\right]\frac{1}{(X_{I.}X_{.1})^{1/2}} & \left[X_{I2} - \left(\frac{X_{I.}X_{.2}}{X_{..}}\right)\right]\frac{1}{(X_{I.}X_{.2})^{1/2}} & \cdots & \left[X_{IJ} - \left(\frac{X_{I.}X_{.J}}{X_{..}}\right)\right]\frac{1}{(X_{I.}X_{.J})^{1/2}} \end{bmatrix}$$

- Dessa forma:

$$Wij = \left[X_{ij} - \frac{X_{i.}X_{.j}}{X_{..}}\right]\frac{1}{(X_{i.}X_{.j})^{1/2}}$$

- Além disso, note que

$$\sum_{i=1}^{I}\left(\frac{X_{ij}}{X_{i.}}\frac{X_{i.}}{X_{..}}\right) = \frac{X_{.j}}{X_{..}}$$

- Seja $\mathbf{A}_{(I \times J)}$, então podemos escrever :

$$\mathbf{A} = \mathbf{U}\Lambda\mathbf{V}'$$

em que
$\mathbf{U}$: colunas formadas pelos autovetores (ortonormalizados) de $\mathbf{AA}'$.
$\mathbf{V}$: coluna formadas pelos autovetores (ortonormalizados) de $\mathbf{A}'\mathbf{A}$.
Λ : matriz diagonal com sua diagonal dada por $(\lambda_1, \lambda_2, ...\lambda_{min(I,J)})'$, que são as raízes quadradas positivas dos autovalores maiores que zero, obtidos a partir das matrizes $\mathbf{AA}'$ ou $\mathbf{A}'\mathbf{A}$ (os autovalores maiores que zero são iguais).

- Decomposição do valor singular de $\mathbf{W}$:

$$\mathbf{W} = \mathbf{U}_{W(I\times I)}\mathbf{\Lambda}_{W(I\times J)}\mathbf{V}'_{W(J\times J)}$$

- $\mathbf{U}_W$ autovetores de $\mathbf{W}\mathbf{W}'$ e $\mathbf{V}_W$ autovetores de $\mathbf{W}'\mathbf{W}$. $\mathbf{\Lambda}_W$ valores singulares de $\mathbf{W}\mathbf{W}'$ ou $\mathbf{W}'\mathbf{W}$, $\lambda_i \geq 0$.
- $\left(\mathbf{R} - \mathbf{1}\mathbf{P}'_c\right) = \mathbf{D}_r^{-1/2}\mathbf{U}_W\mathbf{\Lambda}_W\mathbf{V}'_W\mathbf{D}_c^{1/2}$.
- Resultado: A matriz $\mathbf{A}^*$ que minimiza $||\mathbf{A}^* - (\mathbf{R} - \mathbf{1}\mathbf{P}'_c)||$ é dada por $\mathbf{A}^* = \mathbf{D}_r^{-1/2}\mathbf{U}_W^{(2)}\mathbf{\Lambda}_W^{(2)}\mathbf{V}_W^{\prime(2)}\mathbf{D}_c^{1/2}$ (considerando-se dois autovalores e autovetores).
- Adicionalmente, pode-se provar que $\mathbf{F} = \mathbf{D}_{r(I\times I)}^{-1/2}\mathbf{U}_{W(I\times I)}\mathbf{\Lambda}_{W(I\times J)}$ definem coordenadas relacionadas aos perfis das linhas.

- Podemos demonstrar que $\mathbf{Z} = \mathbf{W}'$ e, de forma análoga, teremos que $\left(\mathbf{C} - \mathbf{1}\mathbf{P}_r'\right) = \mathbf{D}_c^{-1/2}\mathbf{U}_Z\mathbf{\Lambda}_Z\mathbf{V}_Z'\mathbf{D}_r^{1/2}$ (exercício). $\mathbf{Z} = \mathbf{U}_{Z(J\times J)}\mathbf{\Lambda}_{Z(J\times I)}\mathbf{V}'_{Z(I\times I)}$ é a decomposição do valor singular de $\mathbf{Z}$.
- Adicionalmente, pode-se provar que $\mathbf{G} = \mathbf{D}_{c(J\times J)}^{-1/2}\mathbf{U}_{Z(J\times J)}\mathbf{\Lambda}_{Z(J\times I)}$ definem coordenadas relacionadas aos perfis das colunas.
- Calcular e plotar num gráfico de dispersão:

$$
\begin{aligned}
\mathbf{F}^{(2)} &= \mathbf{D}_{r(I\times I)}^{-1/2}\mathbf{U}_{W(I\times 2)}^{(2)}\mathbf{\Lambda}_{W(2\times 2)}^{(2)} \\
\mathbf{G}^{(2)} &= \mathbf{D}_{c(J\times J)}^{-1/2}\mathbf{U}_{Z(J\times 2)}^{(2)}\mathbf{\Lambda}_{Z(2\times 2)}^{(2)}
\end{aligned}
$$

- Note, no entanto, que .

$$
\begin{aligned}
\left(\mathbf{R}-\mathbf{1}\mathbf{P}_c'\right) &= \mathbf{D}_r^{-1/2}\mathbf{U}_W\mathbf{\Lambda}_W\mathbf{V}_W'\mathbf{D}_c^{1/2}\\
\mathbf{D}_r\left(\mathbf{R}-\mathbf{1}\mathbf{P}_c'\right) &= \mathbf{D}_r^{1/2}\mathbf{U}_W\mathbf{\Lambda}_W\mathbf{V}_W'\mathbf{D}_c^{1/2}\\
\mathbf{D}_r\mathbf{R}-\mathbf{D}_r\mathbf{1}\mathbf{P}_c' &= \mathbf{D}_r^{1/2}\mathbf{U}_W\mathbf{\Lambda}_W\mathbf{V}_W'\mathbf{D}_c^{1/2}\\
\mathbf{P}-\mathbf{P}_r\mathbf{P}_c' &= \mathbf{D}_r^{1/2}\mathbf{U}_W\mathbf{\Lambda}_W\mathbf{V}_W'\mathbf{D}_c^{1/2}
\end{aligned}
$$

- $\mathbf{P}-\mathbf{P}_r\mathbf{P}_c$: diferença entre probabilidades observadas e esperadas sob independência.

- Analogamente, note que .

$$
\begin{aligned}
\left(\mathbf{C} - \mathbf{1}\mathbf{P}_r'\right) &= \mathbf{D}_c^{-1/2}\mathbf{U}_Z\mathbf{\Lambda}_W\mathbf{V}_Z'\mathbf{D}_r^{1/2} \\
\mathbf{D}_c\left(\mathbf{C} - \mathbf{1}\mathbf{P}_r'\right) &= \mathbf{D}_c^{1/2}\mathbf{U}_Z\mathbf{\Lambda}_W\mathbf{V}_Z'\mathbf{D}_r^{1/2} \\
\mathbf{D}_c\mathbf{C} - \mathbf{D}_c\mathbf{1}\mathbf{P}_r' &= \mathbf{D}_c^{1/2}\mathbf{U}_Z\mathbf{\Lambda}_W\mathbf{V}_Z'\mathbf{D}_r^{1/2} \\
\mathbf{P} - \mathbf{P}_c\mathbf{P}_r' &= \mathbf{D}_r^{1/2}\mathbf{U}_Z\mathbf{\Lambda}_Z\mathbf{V}_Z'\mathbf{D}_c^{1/2}
\end{aligned}
$$

- $\mathbf{P} - \mathbf{P}_r\mathbf{P}_c$: diferença entre probabilidades observadas e esperadas sob independência.
- Em suma, estamos aproximando as "matrizes de distâncias" por suas decomposições de valores singulares bidimensionais.

- Variabilidade explicada.
- Inércia

$$
\begin{aligned}
& tr\left[\mathbf{D}_r^{-1/2}\left(\mathbf{P}-\mathbf{P}_r\mathbf{P}_c'\right)\mathbf{D}_c^{-1/2}\left(\mathbf{D}_r^{-1/2}\left(\mathbf{P}-\mathbf{P}_r\mathbf{P}_c'\right)\mathbf{D}_c^{-1/2}\right)'\right] \\
= \quad & Q/n = \sum_{k=1}^{J-1}\lambda_k^2 \qquad (1)
\end{aligned}
$$

em que $(\lambda_1, \lambda_2, ..., \lambda_{J-1})$ são os valores singulares obtidos a partir da decomposição do valor singular de $\mathbf{D}_r^{-1/2}\left(\mathbf{P}-\mathbf{P}_r\mathbf{P}_c'\right)\mathbf{D}_c^{-1/2}$.

- Para o caso bidimensional, a inércia será dada por: $\lambda_1^2+\lambda_2^2$ e, a proporção de variabilidade explicada por $\frac{\lambda_1^2+\lambda_2^2}{\sum_{k=1}^{J-1}\lambda_k^2}$.

- Análise de correspondência via R.
- Função *corresp* (biblioteca MASS).
- entrada de dados:

	Qual. de som	**Tec.**	**Pot. do som**	**Rec. técnicos**	**Preços**	**Conf. na marca**
Sony	135	140	95	55	40	60
Aiwa	50	115	40	60	5	15
Gradiente	90	55	20	35	40	10
Philips	60	25	35	10	5	30
Sharp	30	20	5	10	10	20

- Comandos:
  - "Carregar" pacote
  - library(MASS)
  - result.corresp $<-$corresp(m.X.table)
  - biplot(corresp(m.X.table, nf $=$ 2))

- Componentes:

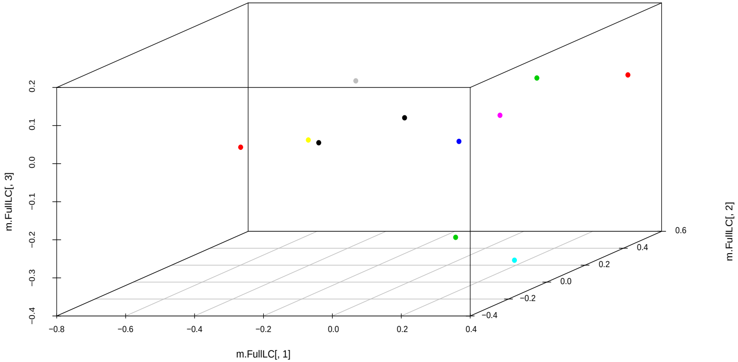
		Componente 1	Componente 2
Perfil das Linhas	Sony	0.032	−0.0804
	Aiwa	−0.4703	−0.0326
	Gradiente	0.1386	0.4188
	Philips	0.3286	−0.3301
	Sharp	0.2989	0.0131
Perfil das Colunas	Qualidade	0.2246	0.0713
	Tecnologia	−0.2924	0.0006
	Potência	−0.002	−0.2773
	Recursos	−0.3368	0.135
	Preço	0.3399	0.532
	Confiança	0.3342	−0.3578

- Inércia

Valor singular	Inércia Princ.	Percentual	Percentual acum.
0.2678	0.0717	52.71	52.71
0.2228	0.0497	36.49	89.20
0.1023	0.0105	7.69	96.88
0.0651	0.0042	3.12	100.00
Total	0.1368	100.00	-

- Calculando inércia via R.
- Calcular a matriz **W** ou a matriz **Z**. Pode-se ser utilizada qualquer uma das matrizes.
- Utilizar o comando `result<- svd(`**W**`)`.
- Extrair os valores singulares através do comando $vs < -result\$d$
- Calcular a inércia tomando-se o quadrado de vs.

- Se utilizissásemos 3 componentes



Utilizando duas