Multidimensional Exploratory Analysis of a Structural Model using a general costructure criterion:

THEME (THematic Equation Model Explorator)

X. Bry — I3M, Univ. Montpellier II
T. Verron — ITG - SEITA, Centre de recherche
P. Redont — I3M, Univ. Montpellier II

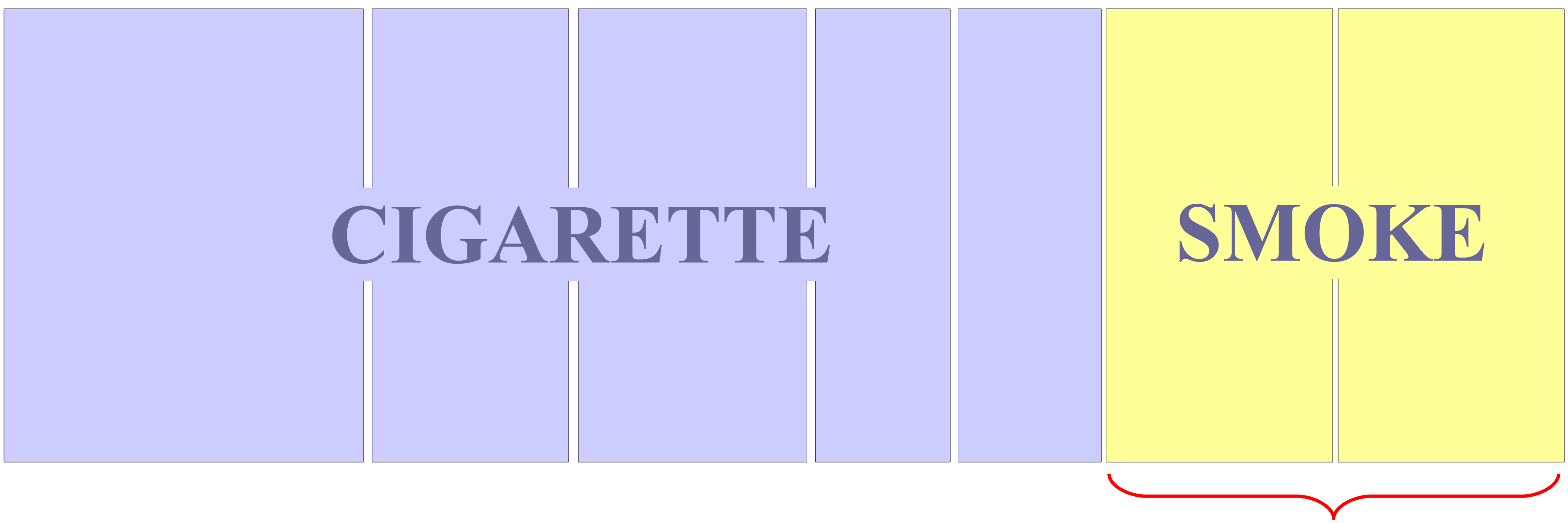
Introducing the Data and Problem:

Data:

52 Variables:

- **15 var.** Tobacco Blend Chemistry
- **5 var.** Paper Combustion
- **8 var.** Tobacco Blend Combustion
- **3 var.** Filtration / ISO smoking
- **3 var.** Filter behaviour / ISO smoking
- **9 var.** Hoffmann smoke contents /Intense smoking
- **9 var.** Hoffmann smoke contents /ISO smoking

19 Observations: **Cigarettes**



Problem: *Regulations → Hoffmann Compounds control* ⇒ **HC modeling**

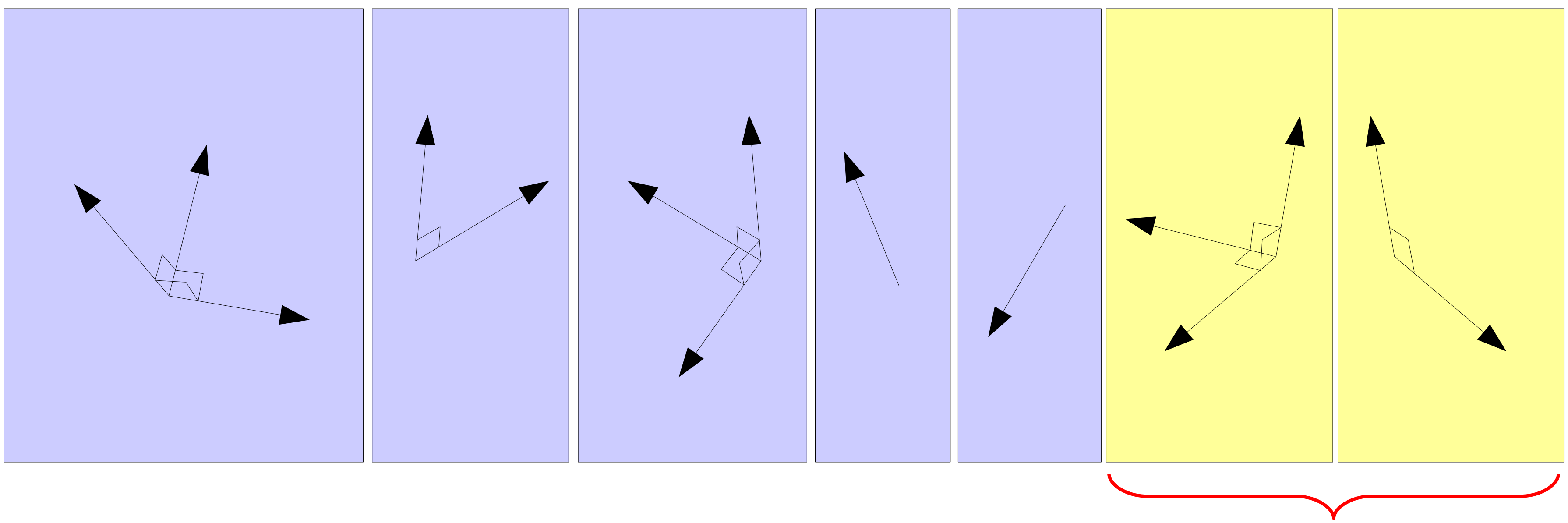
Introducing the Data and Problem:

Data:

52 Variables:

15 var. **Tobacco Blend Chemistry** | **5 var.** **Paper Combustion** | **8 var.** **Tobacco Blend Combustion** | **3 var.** **Filtration / ISO smoking** | **3 var.** **Filter behaviour / ISO smoking** | **9 var.** **Hoffmann smoke contents /Intense smoking** | **9 var.** **Hoffmann smoke contents /ISO smoking**

19 Observations: **Cigarettes**



Problem: *Regulations → Hoffmann Compounds control* ⇒ **HC modeling**

1) The thematic partitioning of variables must be kept (to *separate roles*, and keep *explanatory*)

2) Many (redundant) variables ⇒ Dimension reduction in groups

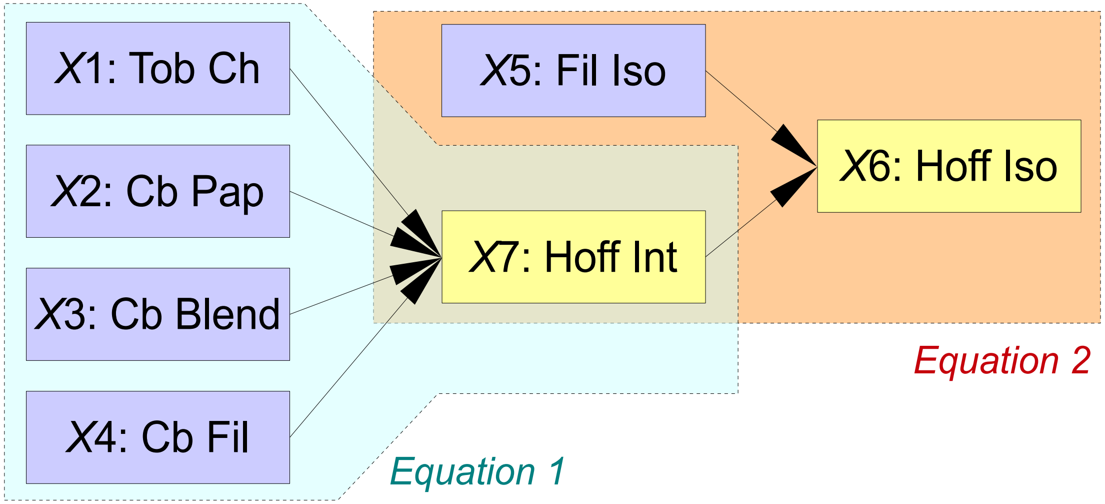
⇒ Look for dimensions: ***reflecting their group's structure***

& interpretable with respect to their theme

Introducing the Data and Problem:

Dependency network of Data:

Thematic (conceptual) model



Model design motivations:

Equation 1:

Hoffmann compounds are generated / transferred to smoke through combustion. Filter only plays a *retention* role (pores blocked in intense mode)

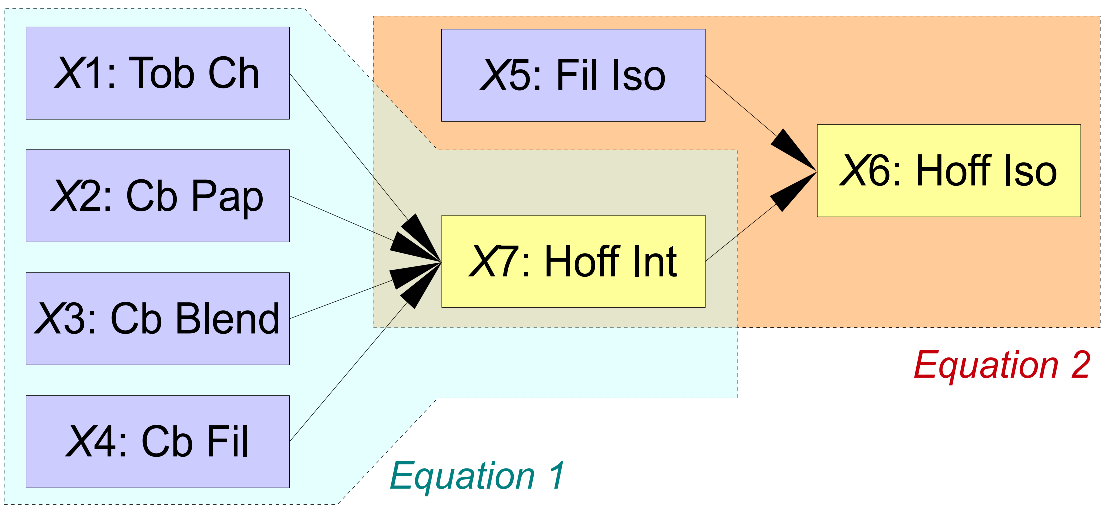
Equation 2:

Final output of Hoffmann compounds is conditioned by other filter properties, as ventilation/dilution.

Introducing the Data and Problem:

Dependency network of Data:

Thematic (conceptual) model



Model design motivations:

Equation 1:

Hoffmann compounds are generated / transferred to smoke through combustion. Filter only plays a *retention* role (pores blocked in intense mode)

Equation 2:

Final output of Hoffmann compounds is conditioned by other filter properties, as ventilation/dilution.

⇒ Structural dimensions should be informative with respect *to the model* too

1) *How many dimensions* do play a proper role? **2) *Which?***

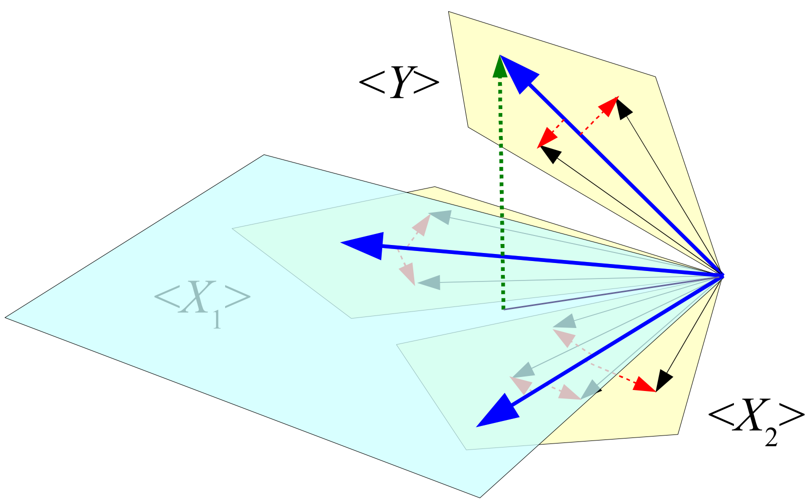
Path modeling methods optimizing a criterion:

- **Likelihood** → **LISREL** (Jöreskog 1975-2002)

- **Residual Sum of Squares** →

  Multiblock Multiway Components and Covariates Regression Models (Smilde, Westerhuis, Bocqué 2000)

  Generalized structured component analysis (Hwang, Takane, 2004).



RSS = RSS(group models)
+ RSS(component-based model)

(minimized via **A**lternated **L**east **S**quares)

- ***Model residuals need weighting: How?***
- ***Convergence problems in case of collinearity (small samples)***
- ***The Methods do not extend PLS Regression to K Predictor Groups.***

based on a covariance criterion...

Extending covariance

- **Multiple Covariance** (Bry, Verron, Cazes 2009)

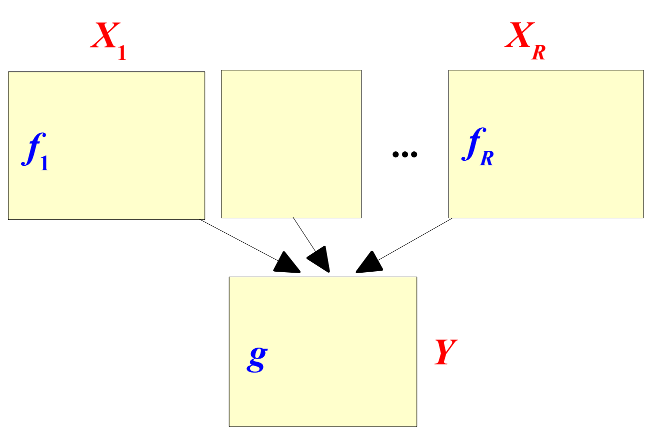
y being linearly modeled as a function of $x^1,\dots,x^S$, *Multiple Covariance of* y *on* $x^1,\dots,x^S$ is:

$$MC(y|x^1,\dots,x^S) = \left[\left(V(y)\prod_{s=1}^{S} V(x^s)\right) R^2(y|x^1,\dots,x^S)\right]^{\frac{1}{2}}$$

Product of all variances — Linear Model Fit

- **Use for single « equation » structural model estimation: SEER** (Bry, Verron, Cazes 2009)

➢ *One component per group:*



$$\max_{\substack{v,\,u_1,\dots,u_R \\ \|v\|^2=1 \\ \forall r,\ \|u_r\|^2=1}} MC^2(Yv|X_1u_1,\dots,X_Ru_R)$$

$$g \mid f_1,\ \dots,\ f_R$$

Extending covariance

- **Multiple Covariance** (Bry, Verron, Cazes 2009)

y being linearly modeled as a function of $x^1, \ldots, x^S$, *Multiple Covariance of* y *on* $x^1, \ldots, x^S$ is:

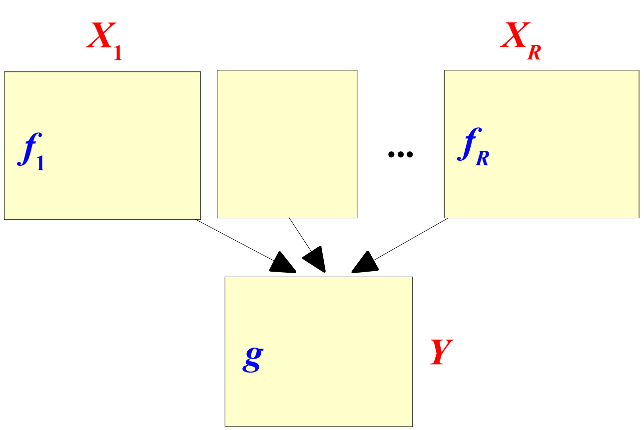
$$MC(y|x^1, \ldots, x^S) = \left[\left(V(y) \prod_{s=1}^{S} V(x^s) \right) R^2(y|x^1, \ldots, x^S) \right]^{\frac{1}{2}}$$

Product of all variances

Linear Model Fit

- **Use for single « equation » structural model estimation: SEER** (Bry, Verron, Cazes 2009)

➢ *One component per group:*



$$\max_{\substack{v, u_1, \ldots, u_R \\ \|v\|^2 = 1 \\ \forall r, \|u_r\|^2 = 1}} MC^2(Yv|X_1 u_1, \ldots, X_R u_R)$$

$$g \mid f_1, \ldots, f_R$$

→ The weighting of Groups is *naturally balanced*

$\nabla \log MC^2 = 0 \quad \Leftrightarrow \quad$ *relative* variations compensate

→ The Method *extends PLS Regression* to K Predictor Groups

Extending covariance

- **Multiple Covariance** (Bry, Verron, Cazes 2009)

y being linearly modeled as a function of $x^1,\ldots,x^S$, *Multiple Covariance of* y *on* $x^1,\ldots,x^S$ is:

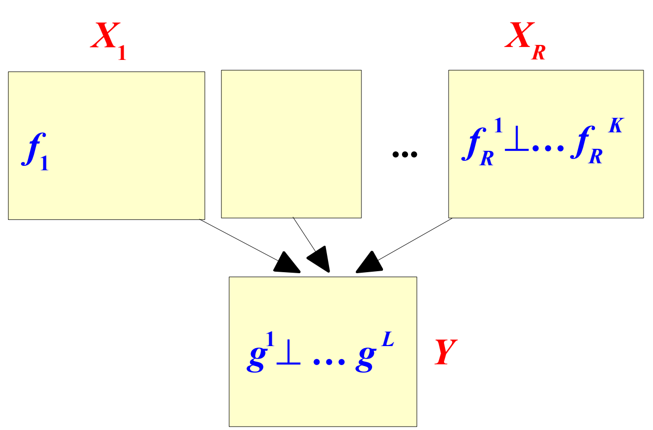
$$MC(y|x^1,\ldots,x^S) = \left[\left(V(y)\prod_{s=1}^{S}V(x^s)\right)R^2(y|x^1,\ldots,x^S)\right]^{\frac{1}{2}}$$

Product of all variances

Linear Model Fit

- **Use for single « equation » structural model estimation: SEER** (Bry, Verron, Cazes 2009)

  - *One component per group:*



$$\max_{\substack{v, u_1, \ldots, u_R \\ \|v\|^2=1 \\ \forall r, \|u_r\|^2=1}} MC^2(Yv|X_1u_1,\ldots,X_Ru_R)$$

$$g \mid f_1, \ldots, f_R$$

→ The weighting of Groups is *naturally balanced*

$\nabla \log MC^2 = 0 \quad \Leftrightarrow$ *relative* variations compensate

→ The Method *extends PLS Regression* to *K* Predictor Groups

  - *Several components per group:* → **Model Local Nesting Principle:** X_r's components $f_r^1, f_r^2\ldots$ are mutually ⊥ and calculated sequentially in one batch, ***controlling for all components in the other groups***

Extending covariance

- **Multiple Covariance** (Bry, Verron, Cazes 2009)

y being linearly modeled as a function of $x^1, \ldots, x^S$, *Multiple Covariance of* y *on* $x^1, \ldots, x^S$ is:

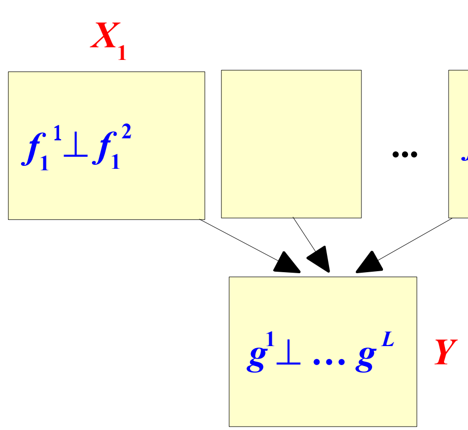
$$MC(y|x^1, \ldots, x^S) = \left[\left(V(y)\prod_{s=1}^{S} V(x^s)\right) R^2(y|x^1, \ldots, x^S)\right]^{\frac{1}{2}}$$

Product of all variances

Linear Model Fit

- **Use for single « equation » structural model estimation: SEER** (Bry, Verron, Cazes 2009)

  - *One component per group:*



$$\max_{\substack{v, u_1, \ldots, u_R \\ \|v\|^2=1 \\ \forall r, \|u_r\|^2=1}} MC^2(Yv|X_1u_1, \ldots, X_Ru_R)$$

$$g \mid f_1, \ldots, f_R$$

→ The weighting of Groups is *naturally balanced*

$\nabla \log MC^2 = 0 \quad \Leftrightarrow \quad$ *relative* variations compensate

→ The Method *extends PLS Regression* to *K* Predictor Groups

  - *Several components per group:* → **Model Local Nesting Principle:** X_r's components $f_r^1, f_r^2 \ldots$ are mutually ⊥ and calculated sequentially in one batch, ***controlling for all components in the other groups***

Extending covariance

- **Multiple Covariance** (Bry, Verron, Cazes 2009)

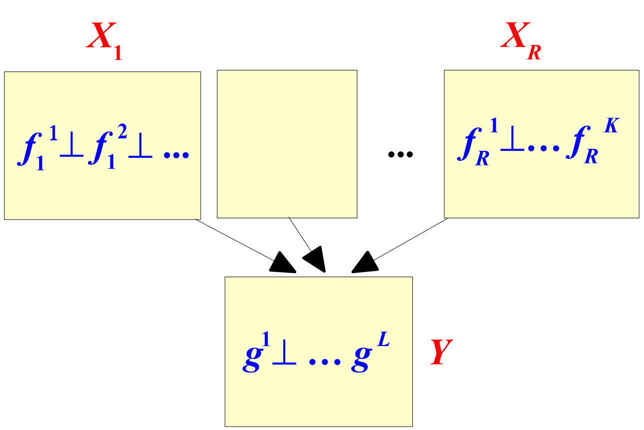
y being linearly modeled as a function of $x^1,\dots,x^S$, *Multiple Covariance of* y *on* $x^1,\dots,x^S$ is:

$$MC(y|x^1,\dots,x^S) = \left[\left(V(y)\prod_{s=1}^{S} V(x^s)\right) R^2(y|x^1,\dots,x^S)\right]^{\frac{1}{2}}$$

Product of all variances — Linear Model Fit

- **Use for single « equation » structural model estimation: SEER** (Bry, Verron, Cazes 2009)

  - *One component per group:*



$$\max_{\substack{v,u_1,\dots,u_R \\ \|v\|^2=1 \\ \forall r, \|u_r\|^2=1}} MC^2(Yv|X_1u_1,\dots,X_Ru_R)$$

$g \mid f_1, \dots, f_R$

→ The weighting of Groups is *naturally balanced*

$\nabla \log MC^2 = 0 \iff$ *relative* variations compensate

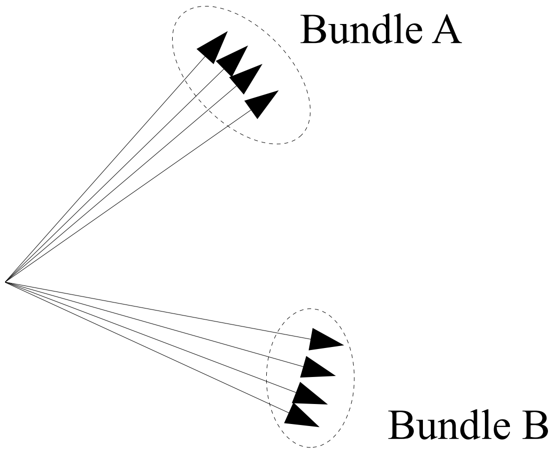
→ The Method *extends PLS Regression* to K Predictor Groups

  - *Several components per group:* → **Model Local Nesting Principle:** X_r's components $f_r^1, f_r^2 \dots$ are mutually ⊥ and calculated sequentially in one batch, ***controlling for all components in the other groups***

Extending covariance

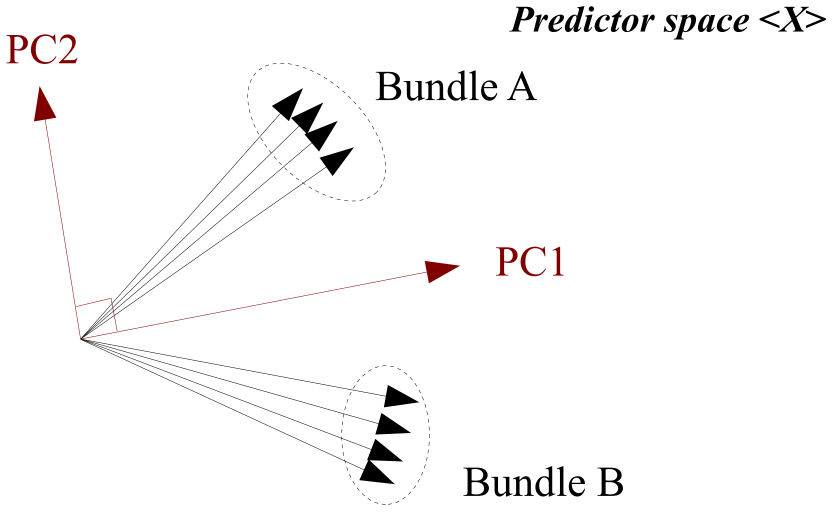
- **Beyond Covariance: Costructure**
  - **Broadened approach to structural strength**

Predictor space <X>



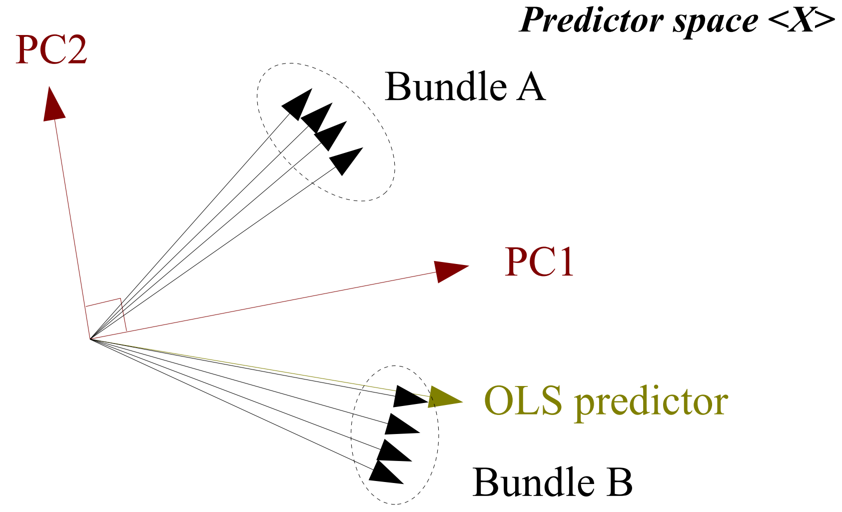
Extending covariance

- **Beyond Covariance: Costructure**
  - **Broadened approach to structural strength**



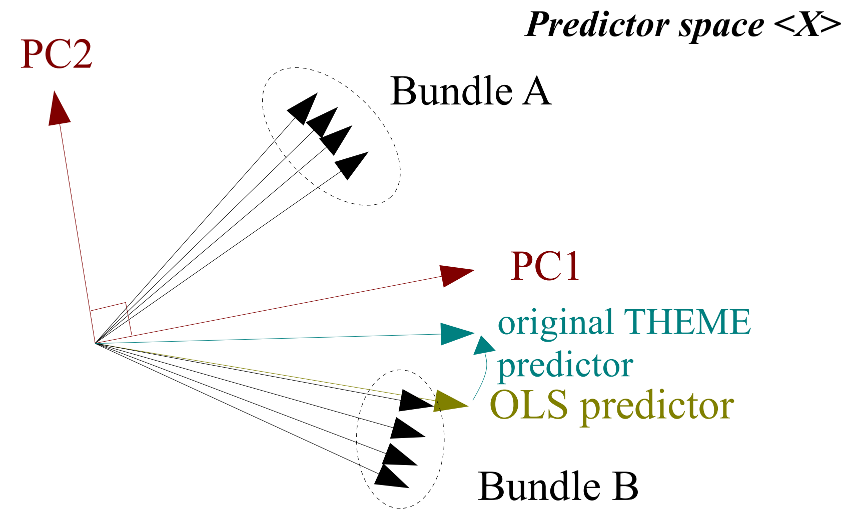
Extending covariance

- **Beyond Covariance: Costructure**
  - **Broadened approach to structural strength**



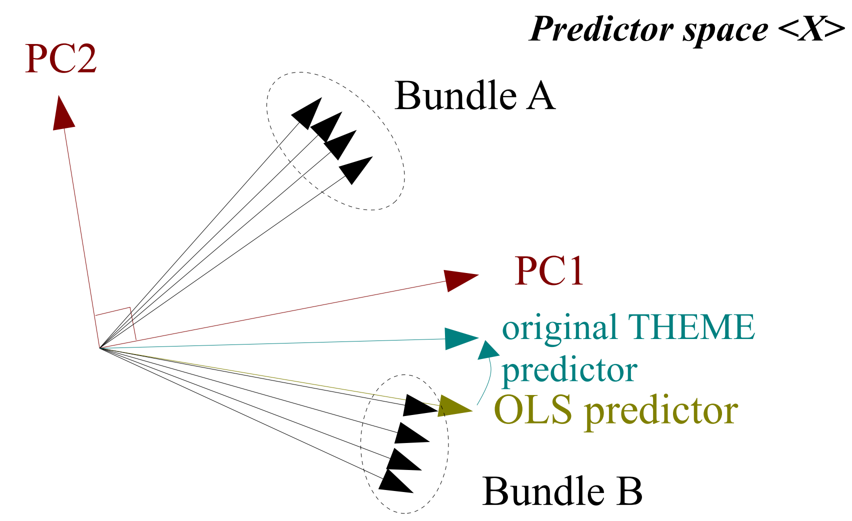
Extending covariance

- **Beyond Covariance: Costructure**
  - **Broadened approach to structural strength**



Extending covariance

- **Beyond Covariance: Costructure**



  - **Broadened approach to structural strength**

  - **General Costructure Criterion**

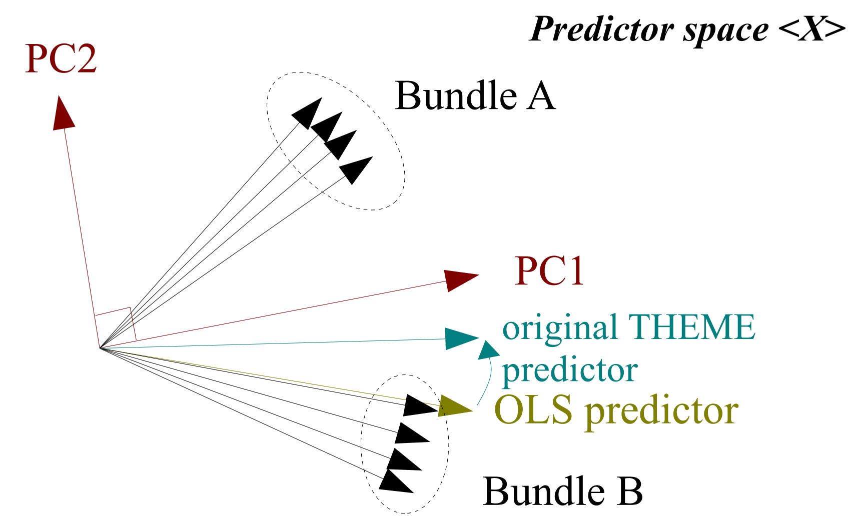
∀ component $f_r = X_r u_r$, $V(f_r) = u_r' X_r' P X_r u_r$ *is* replaced by:

$$S(u_r) = \sum_{h=1,H} (u_r' A_h u_r)^a$$

a = bundle focus parameter

Extending covariance

- **Beyond Covariance: Costructure**



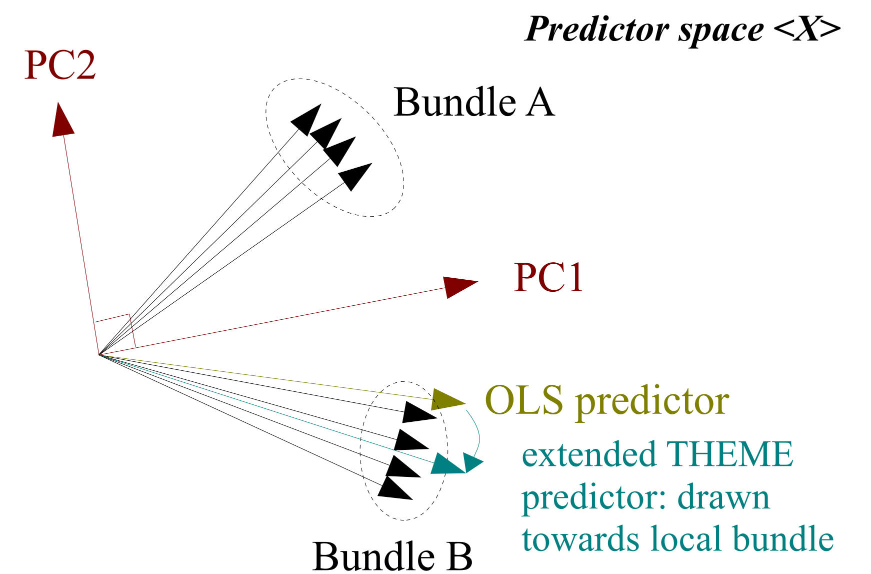
  - **Broadened approach to structural strength**

  - **General Costructure Criterion**

$\forall$ component $f_r = X_r u_r$, $V(f_r) = u_r' X_r' P X_r u_r$ *is* replaced by:

$$S(u_r) = \sum_{h=1,H} (u_r' A_h u_r)^a$$

a = bundle focus parameter



Extending covariance

- **Multiple Co-structure:**

Yv being linearly modeled as a function of $X_1u_1,\ldots,X_Ru_R$, *Multiple Costructure of Yv* on $X_1u_1,\ldots,X_Ru_R$ is:

$$MCS^2(Yv|X_1u_1,\ldots,X_Ru_R) = \left(S(v)\prod_{r=1}^{R}S(u_s)\right)R^2(Yv|X_1u_1,\ldots,X_Ru_R)$$

Product of stuctural strength measures

Linear Model Fit

Extending covariance

- **Multiple Co-structure:**

Yv being linearly modeled as a function of $X_1u_1,\ldots, X_Ru_R$, *Multiple Costructure of* Yv on $X_1u_1,\ldots, X_Ru_R$ is:

$$MCS^2(Yv|X_1u_1,\ldots,X_Ru_R) = \left(S(v)\prod_{r=1}^{R} S(u_s)\right) R^2(Yv|X_1u_1,\ldots,X_Ru_R)$$

Product of stuctural strength measures

Linear Model Fit

- **Extended Multiple Co-structure:**

Let $F = \{f_k = Xu_k;\ k = 1, K\}$ and $G = \{g_j = Yv_j;\ j = 1, J\}$ be two variable groups.
Square *Extended Multiple Costructure of F (powered by γ) and G (powered by δ)* is:

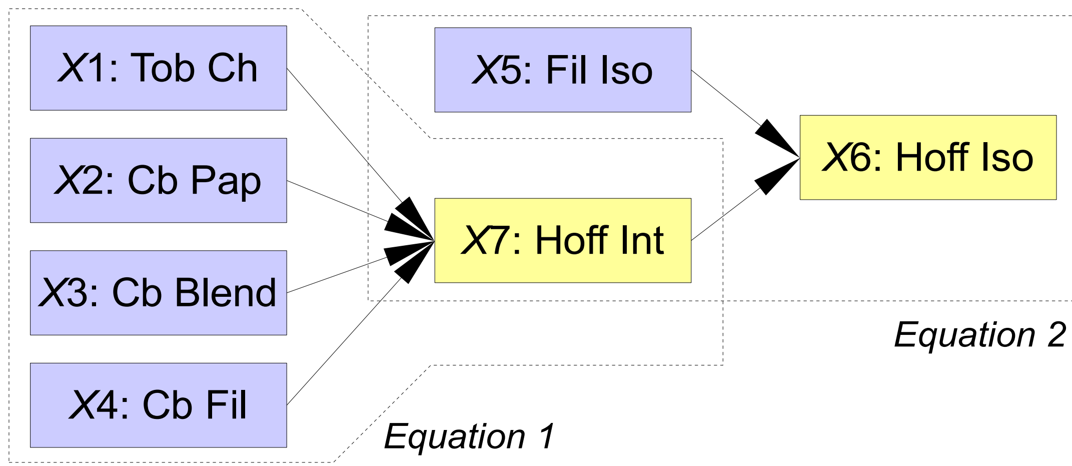
$$EMC^2(F,\gamma;G,\delta)=\left(\prod_{k=1}^{K} S(u_k)\right)^{\gamma}\left(\prod_{j=1}^{J} S(v_j)\right)^{\delta}\frac{\langle \Pi_F|\Pi_G\rangle}{\sqrt{KJ}}$$

Product of stuctural strength measures

Linear Model Fit

Exploring a Multiple Component Equation Model

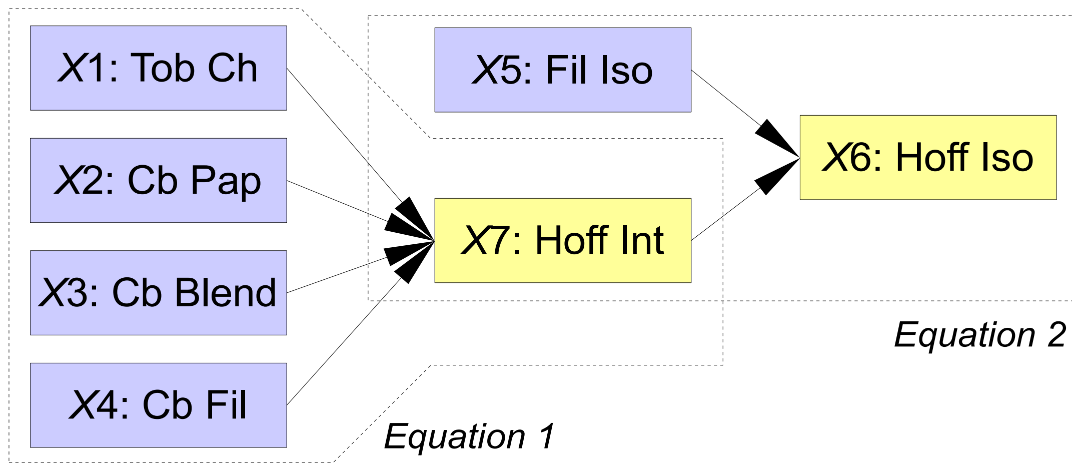
- **System Multiple Covariance Criterion:**



Groups / Equations	Predictive X_1	X_2	X_3	X_4	X_5	X_6	X_7	Dependent X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	×	×	×	×										×
2					×		×						×	

Exploring a Multiple Component Equation Model

- **System Multiple Covariance Criterion:**



Equations \ Groups	Predictive X_1	X_2	X_3	X_4	X_5	X_6	X_7	Dependent X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	×	×	×	×										×
2					×		×						×	

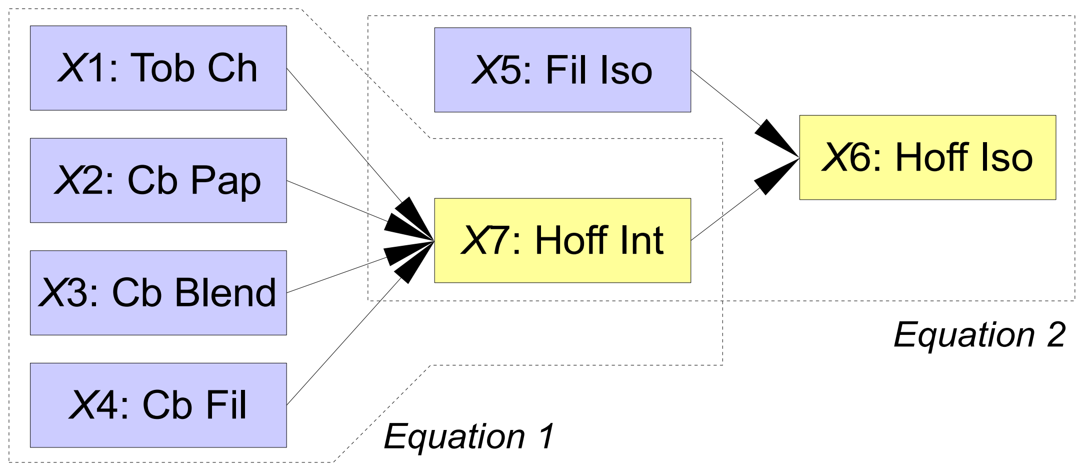
EMC² ($\gamma = \delta = 1$)

$S(u_1)$ … $S(u_4)$ $S(u_7)$ $R^2(X_7u_7 \,|\, X_1u_1, \ldots X_4u_4)$

$S(u_5)$ $S(u_6)$ $S(u_7)$ $R^2(X_6u_6 \,|\, X_5u_5, X_7u_7)$

Exploring a Multiple Component Equation Model

- **System Multiple Covariance Criterion:**



Groups / Equations	Predictive X_1	X_2	X_3	X_4	X_5	X_6	X_7	Dependent X_1	X_2	X_3	X_4	X_5	X_6	X_7
1	×	×	×	×										×
2					×		×						×	

EMC² ($\gamma = \delta = 1$)

$S(u_1)\ \dots\ S(u_4)\qquad S(u_7)\quad R^2(X_7u_7 \mid X_1u_1, \dots X_4u_4)$

$S(u_5)\ S(u_6)\ S(u_7)\quad R^2(X_6u_6 \mid X_5u_5, X_7u_7)$

$S(u_1)\ \dots\ S(u_4)\ S(u_5)\ S(u_6)\ (S(u_7))^2$

$\times R^2(X_7u_7 \mid X_1u_1, \dots X_4u_4)$

$\times R^2(X_6u_6 \mid X_5u_5, X_7u_7)$

$$C = \prod_e EMC^2(\text{Eq. } e) = \prod_{r=1}^{R} \left(S(u_r)\right)^{q_r} \prod_e R^2(\text{Eq. } e)$$

q_r: # of equations involving group X_r

Exploring a Multiple Component Equation Model

- **Maximizing the Global Multiple Covariance Criterion:** $\max_{\substack{u_1, \ldots, u_R \\ \forall r, \|u_r\|^2=1}} C$

C maximized iteratively on each u_r in turn until convergence

$$\Leftrightarrow \max_{u_r \;/\; \|u_r\|^2=1} C(u_r) = \left(S(u_r)\right)^{q_r} \prod_{\text{Eq. } h \text{ involving } X_r} R^2(h)$$

Exploring a Multiple Component Equation Model

- **Maximizing the Global Multiple Covariance Criterion:** $\max_{\substack{u_1, \dots, u_R \\ \forall r, \|u_r\|^2 = 1}} C$

C maximized iteratively on each u_r in turn until convergence

$$\Leftrightarrow \max_{u_r \,/\, \|u_r\|^2 = 1} C(u_r) = \left(S(u_r)\right)^{q_r} \prod_{\text{Eq. } h \text{ involving } X_r} R^2(h)$$

$$S(u_r) = \sum_{h=1,H} (u_r' S_h u_r)^a$$

Exploring a Multiple Component Equation Model

- **Maximizing the Global Multiple Covariance Criterion:** $\max_{\substack{u_1, \ldots, u_R \\ \forall r, \|u_r\|^2=1}} C$

C maximized iteratively on each u_r in turn until convergence

$$\Leftrightarrow \max_{u_r \;/\; \|u_r\|^2=1} C(u_r) = \left(S(u_r)\right)^{q_r} \prod_{\text{Eq. } h \text{ involving } X_r} R^2(h)$$

$$S(u_r) = \sum_{h=1,H} (u_r' S_h u_r)^a$$

X_r dependent

$$R^2(h) = \frac{u_r'(X_r' \Pi_{F_r^h} X_r) u_r}{u_r'(X_r' X_r) u_r}$$

where F^h = components predictive in equation h

Exploring a Multiple Component Equation Model

- **Maximizing the Global Multiple Covariance Criterion:** $\max_{\substack{u_1, \dots, u_R \\ \forall r, \|u_r\|^2=1}} C$

C maximized iteratively on each u_r in turn until convergence

$$\Leftrightarrow \max_{u_r \; / \; \|u_r\|^2=1} C(u_r) = \left(S(u_r)\right)^{q_r} \prod_{\text{Eq. } h \text{ involving } X_r} R^2(h)$$

$$S(u_r) = \sum_{h=1,H} (u_r' S_h u_r)^a$$

X_r dependent

$$R^2(h) = \frac{u_r'(X_r' \Pi_{F_r^h} X_r) u_r}{u_r'(X_r' X_r) u_r}$$

where F^h = components predictive in equation h

X_r predictor of X_d

$$R^2(h) = \frac{u_r'(X_r' A_{rh} X_r) u_r}{u_r'(X_r' B_{rh} X_r) u_r}$$

$$B_{rh} = \Pi_{F^h(-r)^\perp}$$

$$A_{rh} = \frac{1}{\|f_d\|^2}\left[(f_d' \Pi_{F^h(-r)} f_d) B_{rh} + B_{rh}' f_r f_r' B_{rh}\right]$$

Exploring a Multiple Component Equation Model

- **Maximizing the Global Multiple Covariance Criterion:** $\max_{\substack{u_1, \ldots, u_R \\ \forall r, \|u_r\|^2=1}} C$

C maximized iteratively on each u_r in turn until convergence

$$\Leftrightarrow \max_{u_r \,/\, \|u_r\|^2=1} C(u_r) = \left(S(u_r)\right)^{q_r} \prod_{\text{Eq. } h \text{ involving } X_r} R^2(h)$$

$$S(u_r) = \sum_{h=1,H} (u_r' S_h u_r)^a$$

X_r dependent:

$$R^2(h) = \frac{u_r'(X_r' \Pi_{F_r^h} X_r) u_r}{u_r'(X_r' X_r) u_r}$$

where F^h = components predictive in equation h

X_r predictor of X_d:

$$R^2(h) = \frac{u_r'(X_r' A_{rh} X_r) u_r}{u_r'(X_r' B_{rh} X_r) u_r}$$

$$B_{rh} = \Pi_{F^h(-r)^\perp}$$

$$A_{rh} = \frac{1}{\|f_d\|^2}\left[(f_d' \Pi_{F^h(-r)} f_d) B_{rh} + B_{rh}' f_r f_r' B_{rh}\right]$$

$\rightarrow$ Generic form of $C(u_r)$:

$$C(u_r) = \left(\sum_{h=1,H} (u_r' S_h u_r)^a\right)^{\alpha_r} \prod_{l=1}^{q_r} \frac{u_r' T_{rl} u_r}{u_r' W_{rl} u_r}$$

Exploring a Multiple Component Equation Model

- **Generic program** :

$$P: \max_{u_r \,/\, \|u_r\|^2=1} C(u) = \left(\sum_{h=1,H} (u' S_h u)^a \right)^{\alpha} \prod_{l=1}^{q} \frac{u' T_l u}{u' W_l u}$$

Exploring a Multiple Component Equation Model

- **Generic program** :

$$P: \max_{u_r \ / \ \|u_r\|^2=1} C(u) = \left(\sum_{h=1,H} (u' S_h u)^a \right)^{\alpha} \prod_{l=1}^{q} \frac{u' T_l u}{u' W_l u}$$

- **Equivalent unconstrained program** :

$$\boldsymbol{S}: \min_{u \neq 0} \varphi(u) \quad \text{where:} \quad \varphi(u) = \frac{1}{2}[a \alpha u' u - \ln C(v)]$$

→ General minimization software can / should be used

Exploring a Multiple Component Equation Model

- **Generic program :** $P: \max_{u_r \ / \ \|u_r\|^2=1} C(u) = \left(\sum_{h=1,H} (u' S_h u)^a \right)^{\alpha} \prod_{l=1}^{q} \frac{u' T_l u}{u' W_l u}$

- **Equivalent unconstrained program :**

$$\boldsymbol{S}: \min_{u \neq 0} \varphi(u) \quad \text{where:} \quad \varphi(u) = \frac{1}{2}[a \alpha u' u - \ln C(v)]$$

→ General minimization software can / should be used

→ Alternative specific algorithm: $\nabla \varphi(\bar{u}) = 0$

$$\Leftrightarrow \quad \bar{u} = \left[a \alpha I + \sum_{l=1}^{q} \frac{W_l}{\bar{u}' W_l \bar{u}} \right]^{-1} \left[a \alpha \frac{\sum_h (\bar{u}' S_h \bar{u})^{a-1} S_h}{\sum_h (\bar{u}' S_h \bar{u})^a} + \sum_{l=1}^{q} \frac{T_l}{\bar{u}' T_l \bar{u}} \right] \bar{u}$$

suggesting the fixed point algorithm:

$$u(t+1) = \left[a \alpha I + \sum_{l=1}^{q} \frac{W_l}{u(t)' W_l u(t)} \right]^{-1} \left[a \alpha \frac{\sum_h (u(t)' S_h u(t))^{a-1} S_h}{\sum_h (u(t)' S_h u(t))^a} + \sum_{l=1}^{q} \frac{T_l}{u(t)' T_l u(t)} \right] u(t)$$

$$\Leftrightarrow \quad u(t+1) = u(t) - \left[a \alpha I + \sum_{l=1}^{q} \frac{W_l}{u(t)' W_l u(t)} \right]^{-1} \nabla \varphi(u(t)) \qquad (1)$$

Exploring a Multiple Component Equation Model

- **Generic program :** $$P: \max_{u_r \;/\; \|u_r\|^2=1} C(u)=\left(\sum_{h=1,H} (u' S_h u)^a\right)^{\alpha} \prod_{l=1}^{q} \frac{u' T_l u}{u' W_l u}$$

- **Equivalent unconstrained program :**

$$\boldsymbol{S}: \min_{u\neq 0} \varphi(u) \quad \text{where:} \quad \varphi(u)=\frac{1}{2}[a\alpha u'u-\ln C(v)]$$

→ General minimization software can / should be used

→ Alternative specific algorithm: $\nabla\varphi(\bar{u})=0$

$$\Leftrightarrow \quad \bar{u}=\left[a\alpha I+\sum_{l=1}^{q}\frac{W_l}{\bar{u}' W_l \bar{u}}\right]^{-1}\left[a\alpha\frac{\sum_h (\bar{u}' S_h \bar{u})^{a-1} S_h}{\sum_h (\bar{u}' S_h \bar{u})^{a}}+\sum_{l=1}^{q}\frac{T_l}{\bar{u}' T_l \bar{u}}\right]\bar{u}$$

suggesting the fixed point algorithm:

$$u(t+1)=\left[a\alpha I+\sum_{l=1}^{q}\frac{W_l}{u(t)' W_l u(t)}\right]^{-1}\left[a\alpha\frac{\sum_h (u(t)' S_h u(t))^{a-1} S_h}{\sum_h (u(t)' S_h u(t))^{a}}+\sum_{l=1}^{q}\frac{T_l}{u(t)' T_l u(t)}\right]u(t)$$

$$\Leftrightarrow \quad u(t+1)=u(t)-\left[a\alpha I+\sum_{l=1}^{q}\frac{W_l}{u(t)' W_l u(t)}\right]^{-1}\nabla\varphi(u(t)) \qquad (1)$$

descent direction $\boldsymbol{d(t)}$

Exploring a Multiple Component Equation Model

- **Generic program :**

$$P: \max_{u_r \,/\, \|u_r\|^2=1} C(u) = \left(\sum_{h=1,H} (u' S_h u)^a \right)^{\alpha} \prod_{l=1}^{q} \frac{u' T_l u}{u' W_l u}$$

- **Equivalent unconstrained program :**

$$\boldsymbol{S}: \min_{u \neq 0} \varphi(u) \quad \text{where:} \quad \varphi(u) = \frac{1}{2}\left[a \alpha u' u - \ln C(v)\right]$$

→ General minimization software can / should be used

→ Alternative specific algorithm: $\nabla \varphi(\bar{u}) = 0$

$$\Leftrightarrow \quad \bar{u} = \left[a\alpha I + \sum_{l=1}^{q} \frac{W_l}{\bar{u}' W_l \bar{u}} \right]^{-1} \left[a\alpha \frac{\sum_h (\bar{u}' S_h \bar{u})^{a-1} S_h}{\sum_h (\bar{u}' S_h \bar{u})^a} + \sum_{l=1}^{q} \frac{T_l}{\bar{u}' T_l \bar{u}} \right] \bar{u}$$

suggesting the fixed point algorithm:

$$u(t+1) = \left[a\alpha I + \sum_{l=1}^{q} \frac{W_l}{u(t)' W_l u(t)} \right]^{-1} \left[a\alpha \frac{\sum_h (u(t)' S_h u(t))^{a-1} S_h}{\sum_h (u(t)' S_h u(t))^a} + \sum_{l=1}^{q} \frac{T_l}{u(t)' T_l u(t)} \right] u(t)$$

$$\Leftrightarrow \quad u(t+1) = u(t) - \boldsymbol{h(t)} \left[a\alpha I + \sum_{l=1}^{q} \frac{W_l}{u(t)' W_l u(t)} \right]^{-1} \nabla \varphi(u(t)) \qquad (1)$$

descent direction $\boldsymbol{d(t)}$

$\boldsymbol{h(t) = 1}$ works, but using $\boldsymbol{h(t) > 0}$ improves convergence rate.
If chosen according to the Wolfe, or Goldstein-Price, rule: convergence to critical point guaranteed.

- Numerous simulations → **(almost) always *global* minimum**
- (1) **numerically faster** than classical **gradient descent.**

Exploring a Multiple Component Equation Model

- **What if we want several components per group?**
  - $\boldsymbol{K_r}$ **given** ; $\boldsymbol{X_r} \rightarrow \{f_r^1, f_r^2 \ldots f_r^{Kr}\}$ mutually $\perp$

Exploring a Multiple Component Equation Model

- **What if we want several components per group?**
  - $\boldsymbol{K_r}$ **given** ; $\boldsymbol{X_r} \rightarrow \{f_r^1, f_r^2 \ldots f_r^{Kr}\}$ mutually $\perp$

Model Local Nesting Principle:

f_r^1 calculated, *all components in the other groups considered given;*

$\rightarrow \quad X_r^1 = X_r - (1/\| f_r^1 \|) f_r^1 f_r^{1\prime} X_r$ = group of residuals of X_r regressed on f_r^1

f_r^2 calculated with group X_r^1 , *all components in the other groups considered given,* ***plus*** f_r^1 ;

$\rightarrow \quad X_r^2 = X_r^1 - (1/\| f_r^2 \|) f_r^2 f_r^{2\prime} X_r^1$ = group of residuals of X_r regressed on $\{ f_r^1 , f_r^2 \}$

etc.

Exploring a Multiple Component Equation Model

- **What if we want several components per group?**
  - **K_r given** ; X_r → $\{f_r^1, f_r^2 \ldots f_r^{Kr}\}$ mutually ⊥

    Model Local Nesting Principle:

    f_r^1 calculated, *all components in the other groups considered given;*

    → $X_r^1 = X_r - (1/\|f_r^1\|) f_r^1 f_r^{1\prime} X_r$ = group of residuals of X_r regressed on f_r^1

    f_r^2 calculated with group X_r^1, *all components in the other groups considered given,* ***plus*** f_r^1 ;

    → $X_r^2 = X_r^1 - (1/\|f_r^2\|) f_r^2 f_r^{2\prime} X_r^1$ = group of residuals of X_r regressed on $\{f_r^1, f_r^2\}$

    etc.

  - **Finding good K_r values through backward selection:**

    Starting with **large** K_r's → concentrating on "proper" effects

    → K_r's maybe **too large**! (over-fitting, on structurally weak dimensions... up to noise).

Exploring a Multiple Component Equation Model

- **What if we want several components per group?**

  - **K_r given** ; **X_r** $\rightarrow$ $\{f_r^1, f_r^2 \ldots f_r^{Kr}\}$ mutually $\perp$

    Model Local Nesting Principle:

    f_r^1 calculated, *all components in the other groups considered given;*

    $\rightarrow$ $X_r^1 = X_r - (1/\|f_r^1\|) f_r^1 f_r^{1\prime} X_r$ = group of residuals of X_r regressed on f_r^1

    f_r^2 calculated with group X_r^1, *all components in the other groups considered given,* ***plus*** f_r^1 ;

    $\rightarrow$ $X_r^2 = X_r^1 - (1/\|f_r^2\|) f_r^2 f_r^{2\prime} X_r^1$ = group of residuals of X_r regressed on $\{ f_r^1, f_r^2 \}$

    etc.

  - **Finding good K_r values through backward selection:**

    Starting with **large** K_r's $\rightarrow$ concentrating on "proper" effects

    $\rightarrow$ K_r's maybe **too large**! (over-fitting, on structurally weak dimensions... up to noise).

$\rightarrow$ Problem: given estimated model with $(K_1, \ldots, K_R)$ components:
which of the K_r-rank components could / should we *preferably* remove?

i.e. with the smallest possible drop in...

- predictive power? Cross-validation error-rate
- explanatory power? Interpretability
- the global criterion? "technically" handy

Numeric experiments

Experiments:

Parameter values: $a = 2$, $\alpha = q = 2$;
Size 100×100 s.d.p. matrices with various eigenvalues patterns , 50 times, with 50 starting points.

$\rightarrow$ There are local maxima, but a seemingly global maximum is reached most of the time.

Numeric experiments

Experiments:

Parameter values: $a = 2$, $\alpha = q = 2$;
Size 100×100 s.d.p. matrices with various eigenvalues patterns , 50 times, with 50 starting points.

$\rightarrow$ There are local maxima, but a seemingly global maximum is reached most of the time.

Compared performance of the three maximization methods

(1) ***Standard maximization subroutines*** ...
- demand gradient threshold not too low (flat limit of function makes the routine oversensitive to calculus error noise)

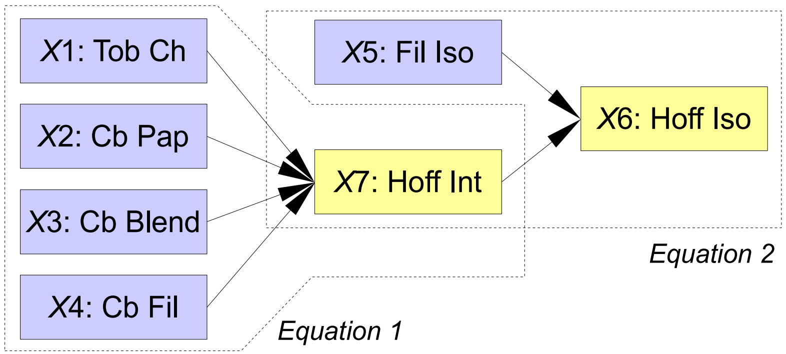
(2) ***Fixed point algorithm*** ($h = 1$): no problem encountered ;
- may reach arbitrary low gradient;
- 2 to 3 times slower than (1).

Slower, but more Robust

(3) ***h* optimized through Wolfe rule**:
- theoretical safeguard... useless in practice;
- demanding a too low gradient results in instability in certain cases.

Application to cigarette data

Multiple Covariance criterion



- Initially: $K = 3$ components *per* group
- Remove rank K_r component alternately in each (predictor) group X_r

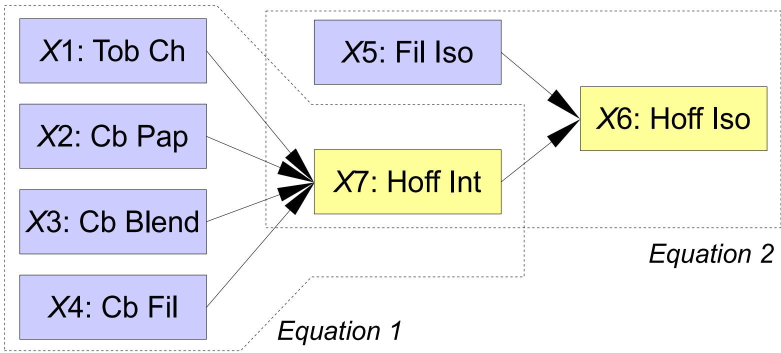
  → 6 « shrunk » models → Evaluated *via* cross-validation

  → Best model selected.
- Resume with selected model

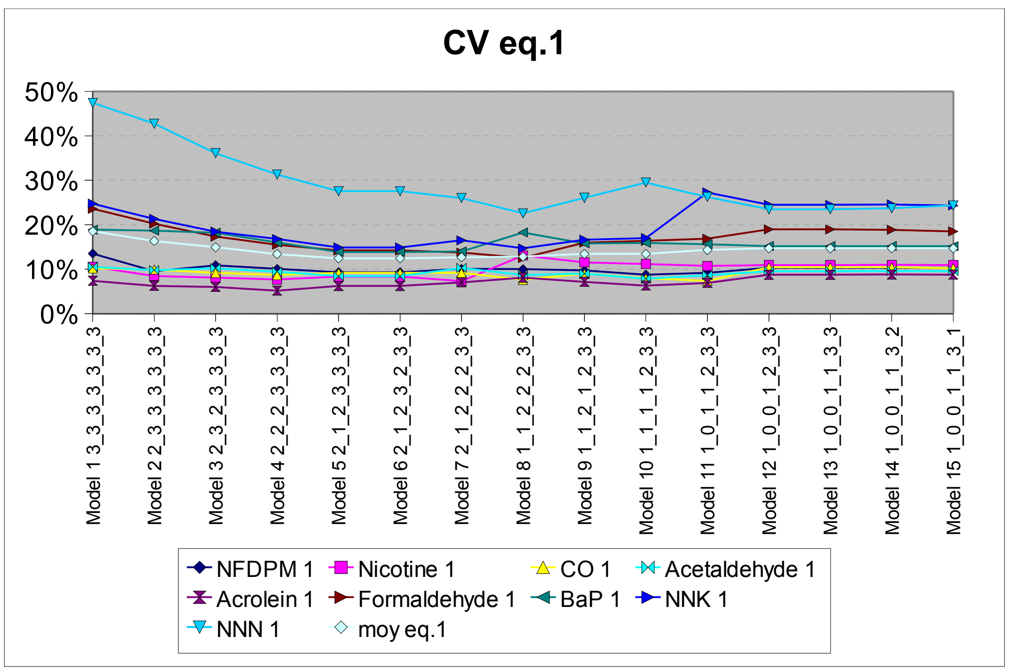
Triple sample:
- Calibration
- Test & selection
- Validation

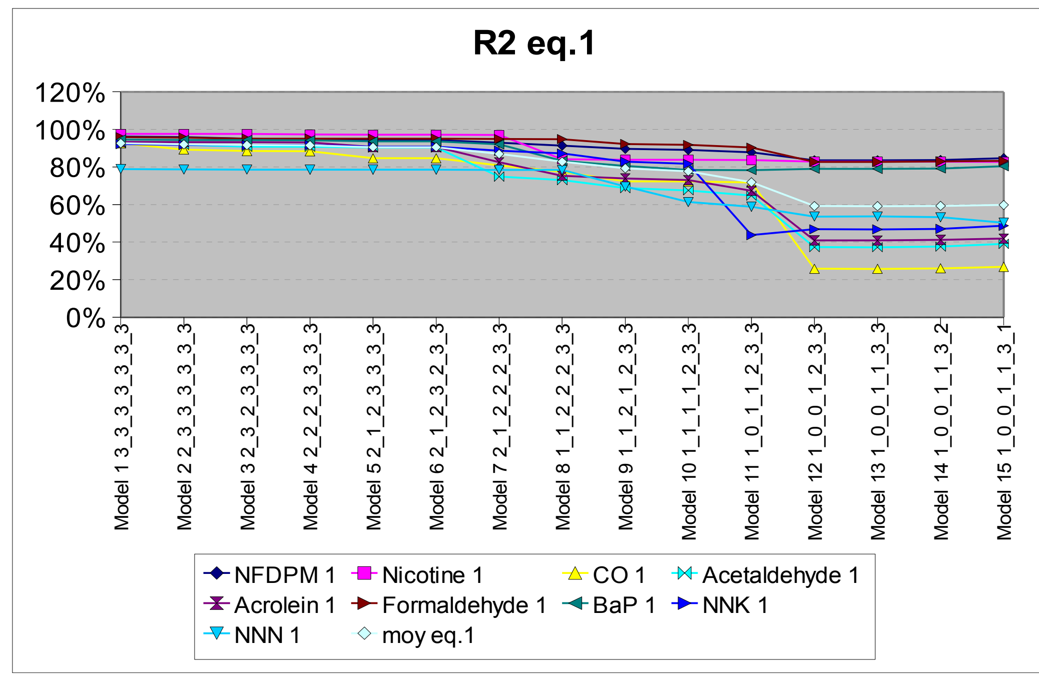
Application to cigarette data

Multiple Covariance criterion



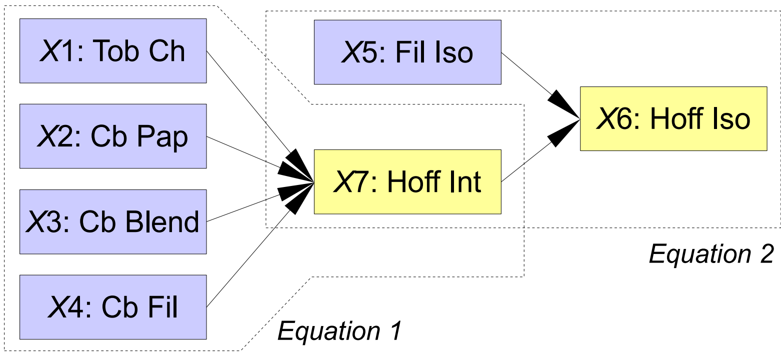
- Initially: $K = 3$ components *per* group
- Remove rank K_r component alternately in each (predictor) group X_r
  → 6 « shrunk » models → Evaluated *via* cross-validation
  → Best model selected.
- Resume with selected model





Application to cigarette data

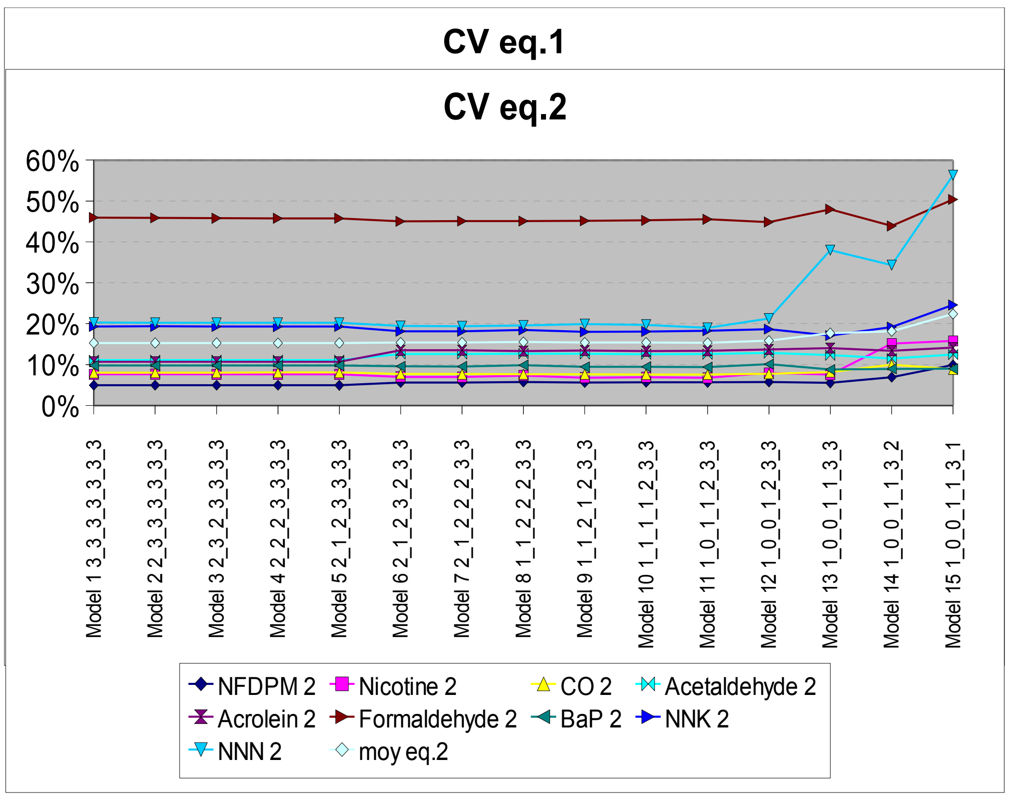
Multiple Covariance criterion

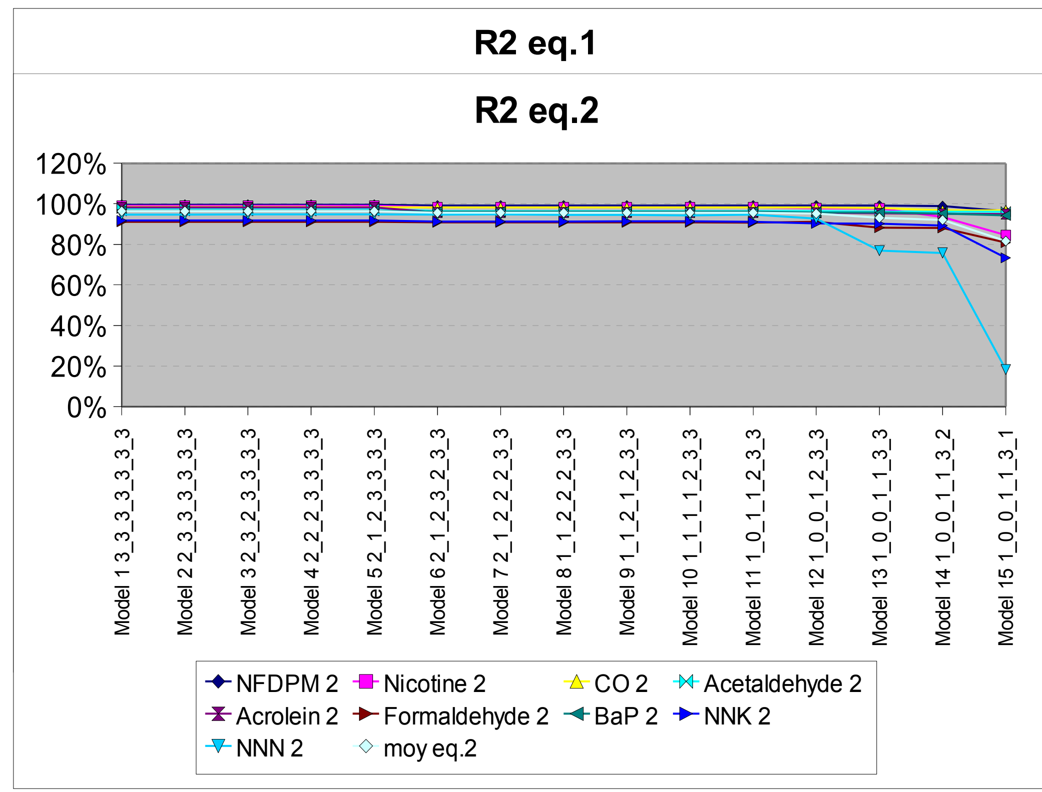


- Initially: $K = 3$ components *per* group
- Remove rank K_r component alternately in each (predictor) group X_r

  → 6 « shrunk » models → Evaluated *via* cross-validation

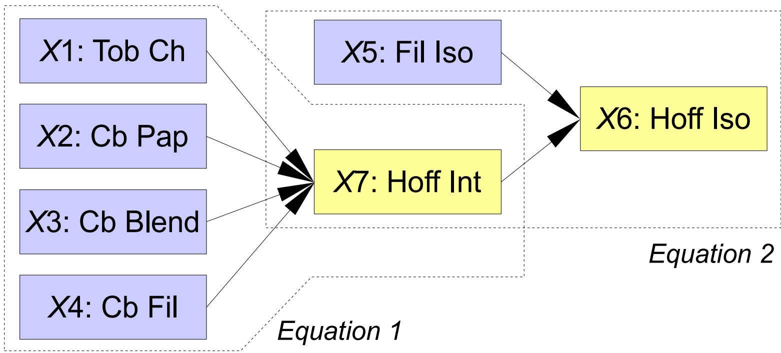
  → Best model selected.
- Resume with selected model



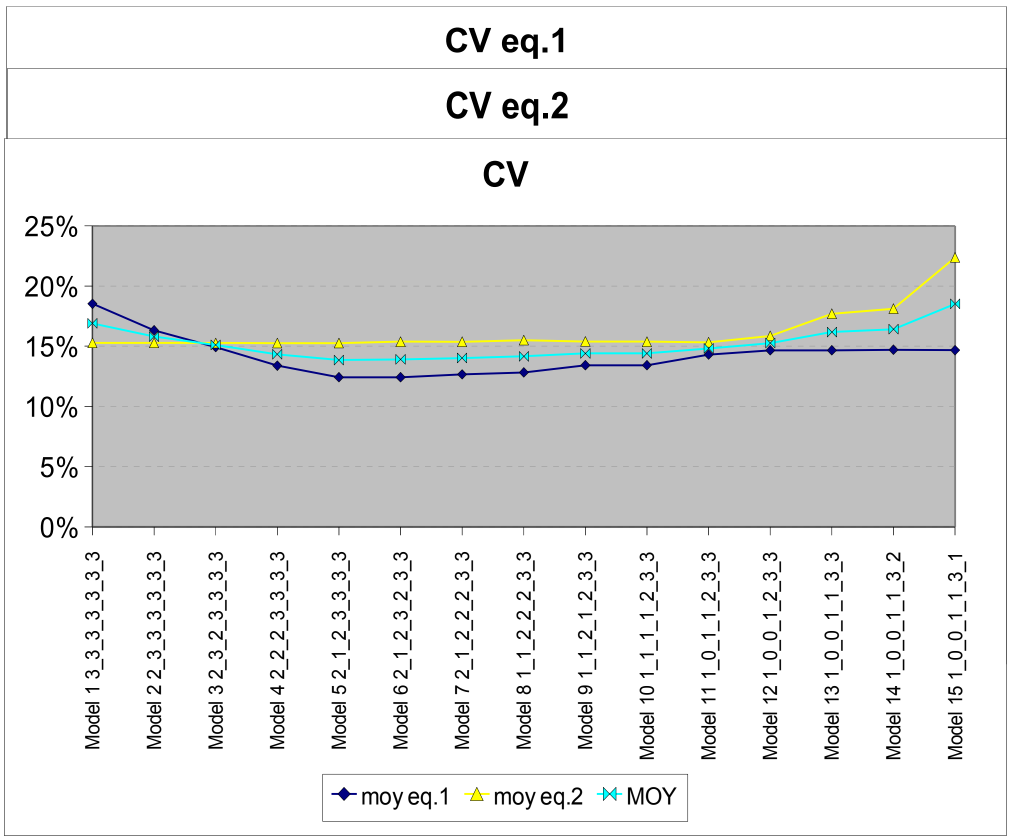


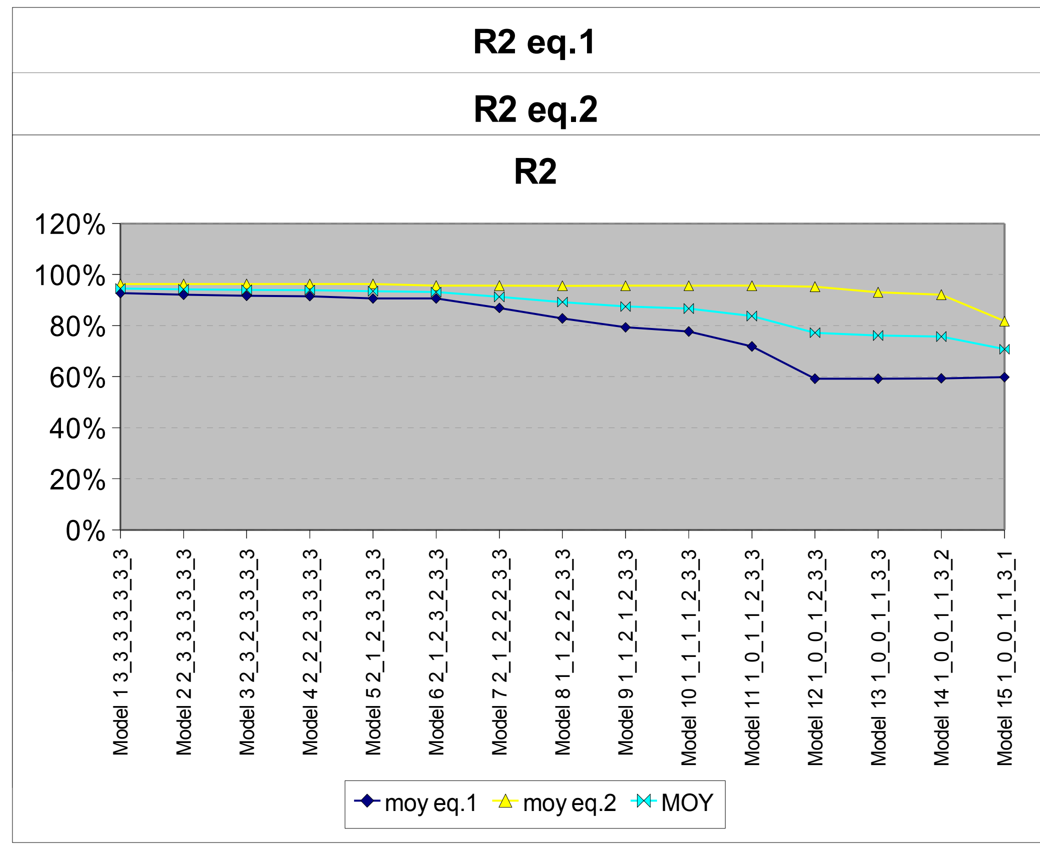
Application to cigarette data

Multiple Covariance criterion



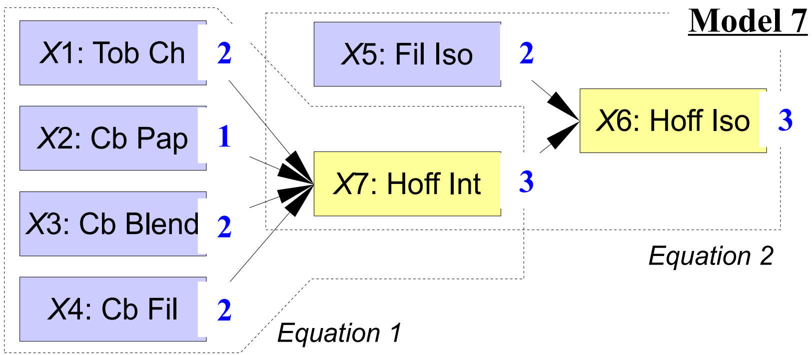
- Initially: $K = 3$ components *per* group
- Remove rank K_r component alternately in each (predictor) group X_r
  → 6 « shrunk » models → Evaluated *via* cross-validation
  → Best model selected.
- Resume with selected model



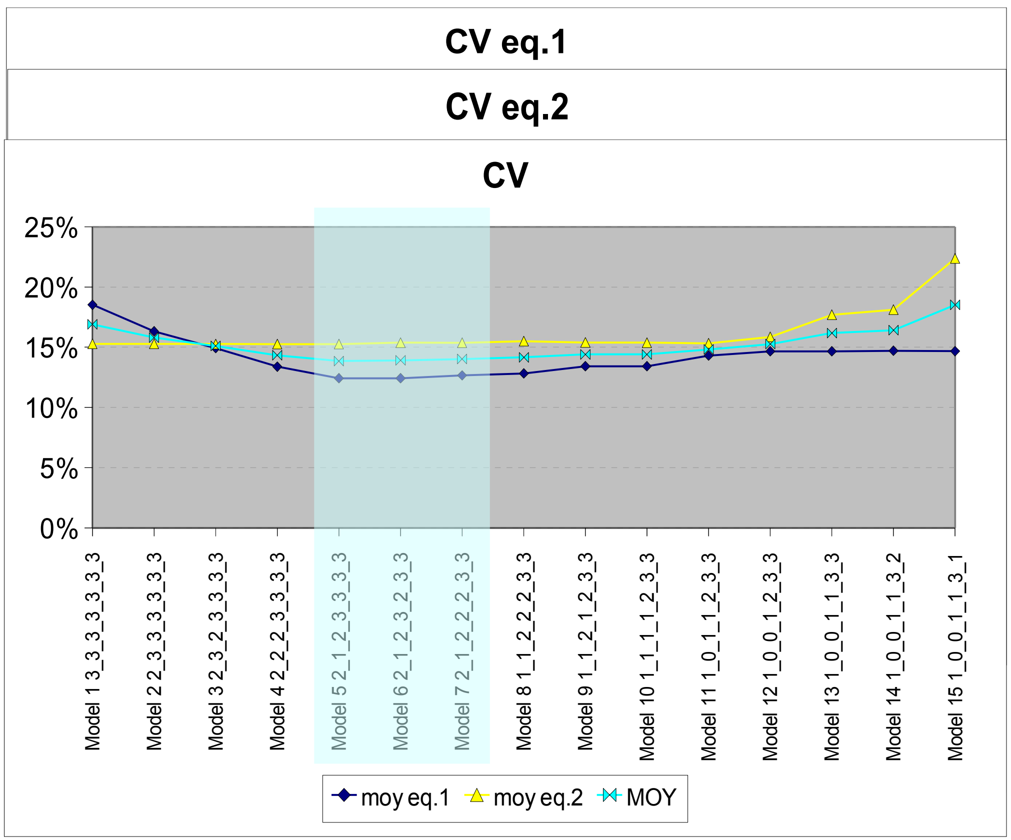


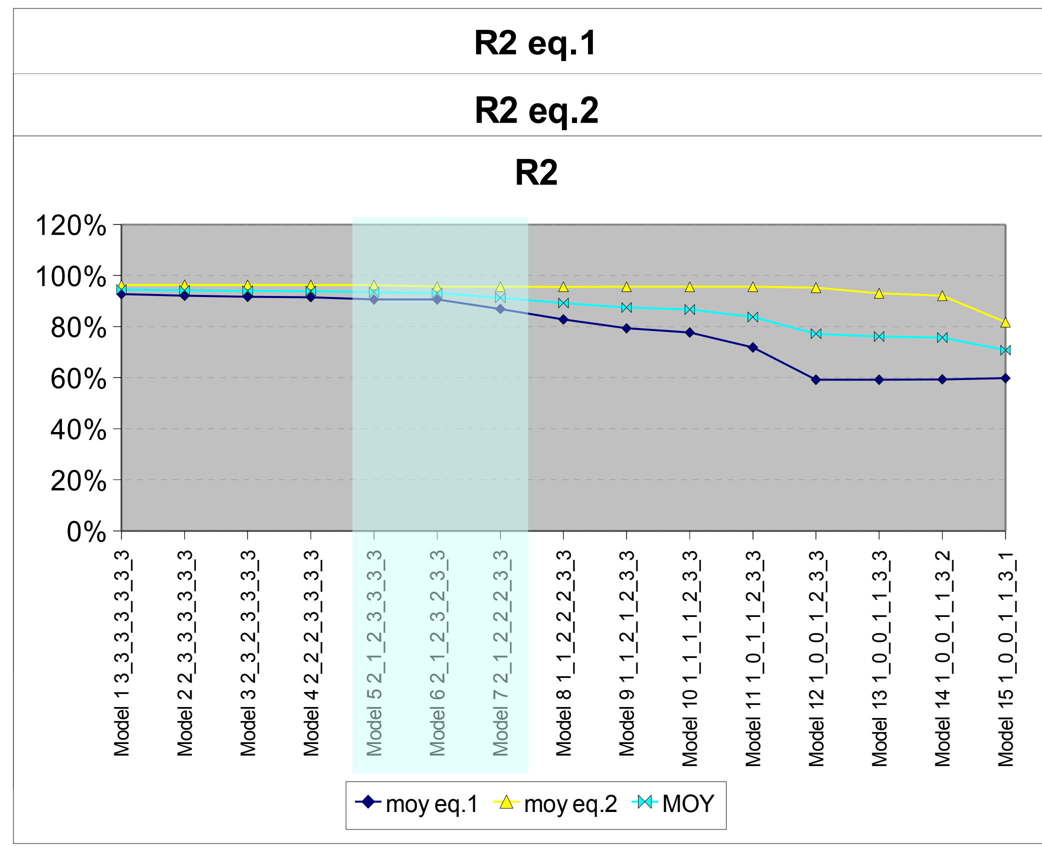
Application to cigarette data

Multiple Covariance criterion



- Initially: $K = 3$ components *per* group
- Remove rank K_r component alternately in each (predictor) group X_r
  → 6 « shrunk » models → Evaluated *via* cross-validation
  → Best model selected.
- Resume with selected model

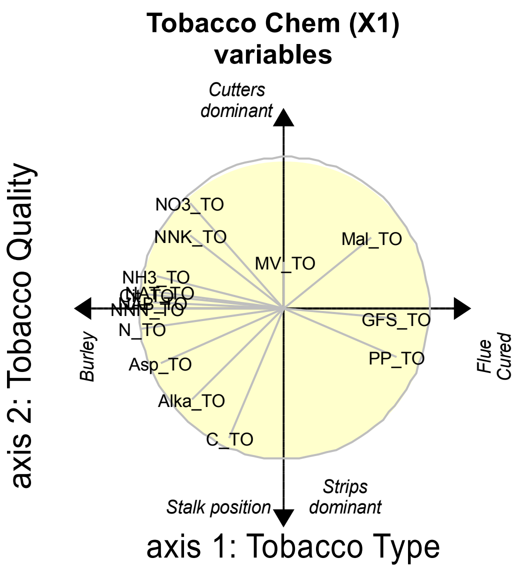


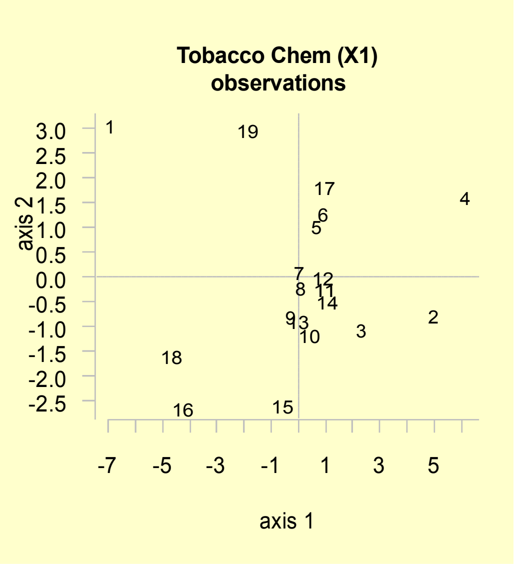


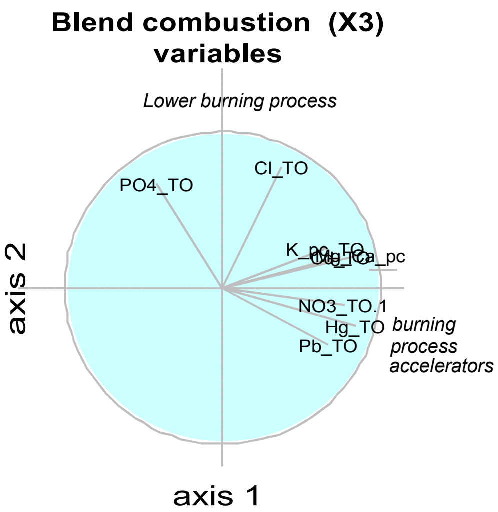
→ **Models 5, 6, 7**

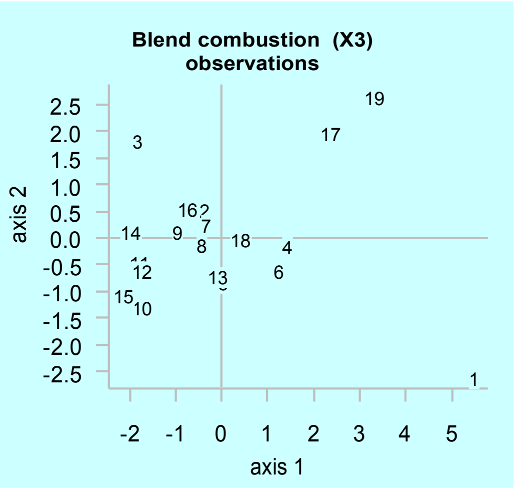
Application to cigarette data

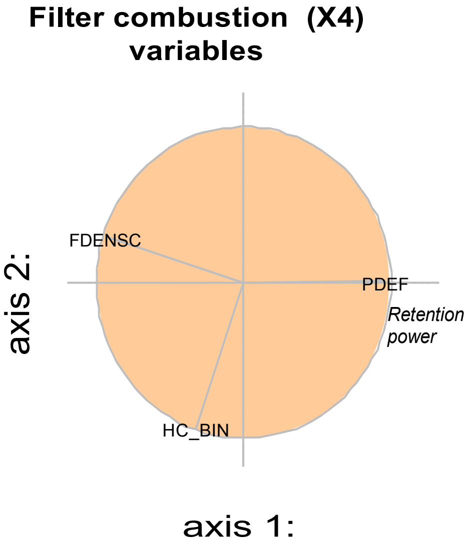
Component-planes for exogenous groups (model 7)

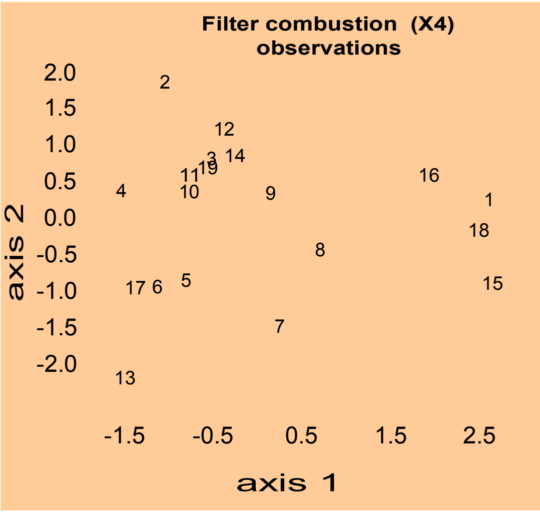


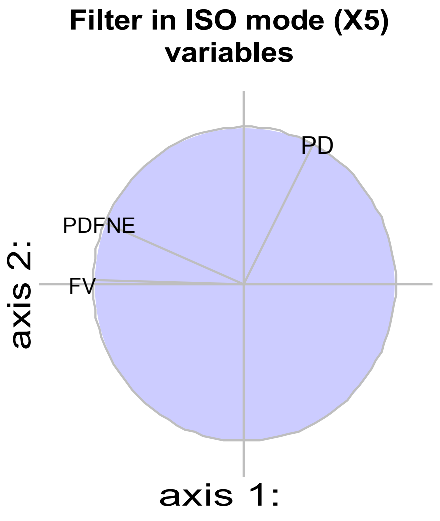


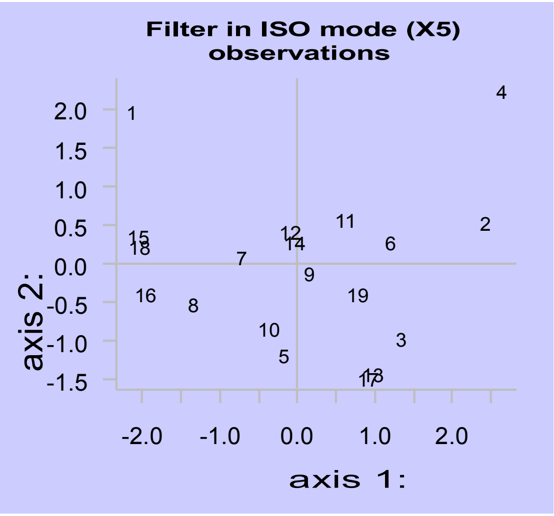






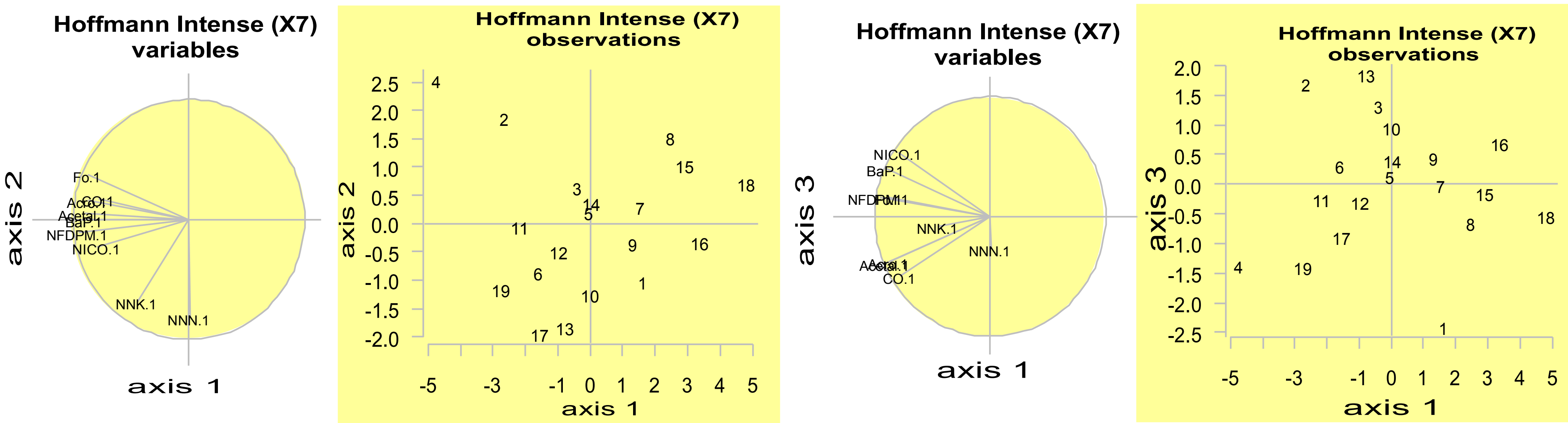


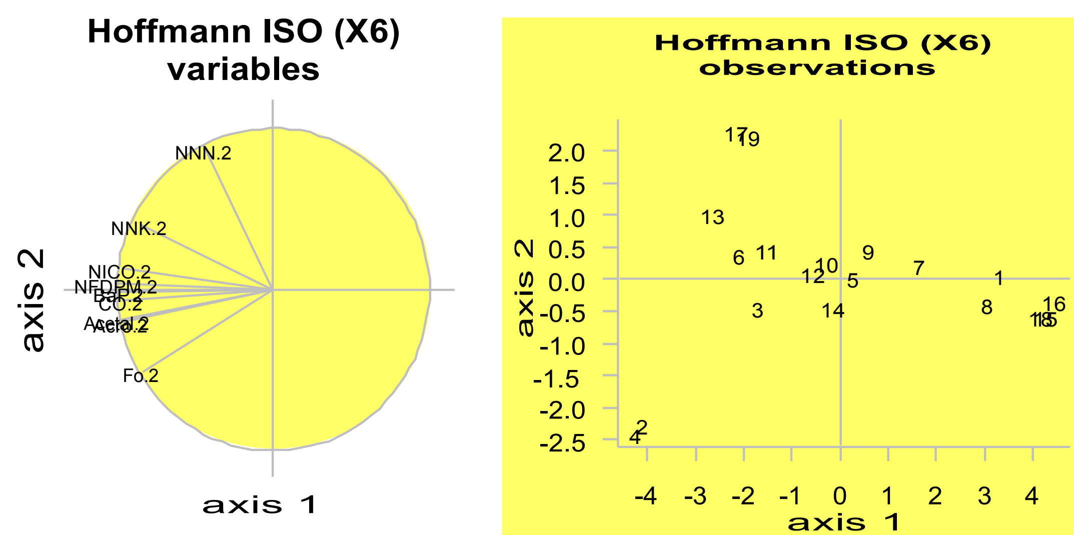




Application to cigarette data

Component-planes for dependent groups (model 7)





- Roughly similar structures of predicted Hoffmann compounds in Intense and ISO modes.
- Positive correlation of all compounds reflects the filter ventilation effect.
- NNK and NNN are strongly related to tobacco type.

Application to cigarette data

Hoff. Compounds regressed on model 7 Components

Equation 1

.	*	**	***

		NFDPM	Nicotine	CO	Acetaldehyde	Acrolein	Formaldehyde	BaP	NNK	NNN
Group 1	F1	0,03	-0,09	0,24	0,13	0,21	0,28	0,02	-0,40	-0,32
	F2	-0,22	-0,64	0,34	0,26	0,48	0,00	-0,53	-0,21	0,06
Group 2	F1	-0,19	-0,28	0,09	-0,06	-0,06	-0,10	-0,27	-0,47	-0,07
Group 3	F1	0,30	0,40	0,16	0,13	-0,03	0,17	0,41	0,19	0,05
	F2	0,06	0,06	-0,12	0,02	0,02	0,03	0,15	-0,18	0,38
Group 4	F1	-0,67	-1,02	0,10	-0,12	0,11	-0,09	-0,74	-0,95	-0,46
	F2	0,17	0,10	0,24	0,22	0,10	0,18	0,23	0,25	-0,34

Equation 2

		NFDPM	Nicotine	CO	Acetaldehyde	Acrolein	Formaldehyde	BaP	NNK	NNN
Group 7	F1	-0,13	-0,13	-0,08	-0,11	-0,10	-0,04	-0,22	-0,38	0,13
	F2	-0,12	-0,20	0,01	0,02	0,02	0,17	-0,07	-0,37	-0,48
	F3	0,06	0,22	-0,15	0,06	0,13	0,18	0,12	0,14	-0,60
Group 5	F1	0,50	0,43	0,60	0,50	0,51	0,51	0,33	-0,04	0,61
	F2	-0,01	-0,05	-0,04	0,08	0,08	0,25	0,00	0,01	-0,57

Application to cigarette data

Hoff. Compounds regressed on model 7 Components

Equation 1

. * ** ***

		NFDPM	Nicotine	CO	Acetaldehyde	Acrolein	Formaldehyde	BaP	NNK	NNN
Group 1	F1	0,03	-0,09	0,24	0,13	0,21	0,28	0,02	-0,40	-0,32
	F2	-0,22	-0,64	0,34	0,26	0,48	0,00	-0,53	-0,21	0,06
Group 2	F1	-0,19	-0,28	0,09	-0,06	-0,06	-0,10	-0,27	-0,47	-0,07
Group 3	F1	0,30	0,40	0,16	0,13	-0,03	0,17	0,41	0,19	0,05
	F2	0,06	0,06	-0,12	0,02	0,02	0,03	0,15	-0,18	0,38
Group 4	F1	-0,67	-1,02	0,10	-0,12	0,11	-0,09	-0,74	-0,95	-0,46
	F2	0,17	0,10	0,24	0,22	0,10	0,18	0,23	0,25	-0,34

Equation 2

		NFDPM	Nicotine	CO	Acetaldehyde	Acrolein	Formaldehyde	BaP	NNK	NNN
Group 7	F1	-0,13	-0,13	-0,08	-0,11	-0,10	-0,04	-0,22	-0,38	0,13
	F2	-0,12	-0,20	0,01	0,02	0,02	0,17	-0,07	-0,37	-0,48
	F3	0,06	0,22	-0,15	0,06	0,13	0,18	0,12	0,14	-0,60
Group 5	F1	0,50	0,43	0,60	0,50	0,51	0,51	0,33	-0,04	0,61
	F2	-0,01	-0,05	-0,04	0,08	0,08	0,25	0,00	0,01	-0,57

Equation 1

. * ** ***

		NFDPM	Nicotine	CO	Acetaldehyde	Acrolein	Formaldehyde	BaP	NNK	NNN
Group 1	F1	0,03	-0,09	0,24	0,13	0,21	0,28	0,02	-0,40	-0,32
	F2	-0,22	-0,64	0,34	0,26	0,48	0,00	-0,53	-0,21	0,06
	C_TO	0,99	0,25	-1,02	-37,51	-6,55	-1,60	1,71	6,58	1,14
	Mal_TO	-0,63	-0,18	0,88	30,60	5,23	3,03	-1,14	-7,07	-7,91
	N_TO	0,19	0,13	-1,11	-33,58	-5,43	-8,17	0,51	12,52	28,28
	PP_TO	0,92	0,16	-0,07	-9,60	-2,10	6,05	1,42	-4,67	-25,84
	MV_TO	0,00	0,00	0,00	0,14	0,02	0,00	-0,01	-0,02	0,01
	Asp_TO	2,50	0,84	-4,91	-162,08	-27,19	-24,08	4,80	45,33	74,36
	Cit_TO	-0,25	-0,01	-0,34	-7,62	-1,04	-4,74	-0,32	5,68	18,09
	NO3_TO	-2,53	-0,53	1,31	58,58	10,86	-7,11	-4,13	-0,82	37,11
	Alka_TO	1,67	0,46	-2,15	-75,58	-13,00	-6,33	2,98	16,32	14,87
	GFS_TO	0,05	0,00	0,09	2,14	0,31	1,10	0,05	-1,36	-4,13
	NH3_TO	-4,76	-0,64	-1,70	-10,28	1,39	-49,24	-6,92	49,83	197,75
	NAB_TO	-3,71	0,52	-12,94	-342,77	-51,81	-138,27	-3,06	181,82	510,65
	NAT_TO	-0,29	-0,01	-0,38	-8,56	-1,17	-5,36	-0,37	6,42	20,47
	NNK_TO	-2,83	-0,56	1,05	53,45	10,24	-11,55	-4,53	4,23	54,31
	NNN_TO	-0,06	0,02	-0,32	-8,88	-1,37	-3,20	-0,02	4,33	11,66
Group 2	F1	-0,19	-0,28	0,09	-0,06	-0,06	-0,10	-0,27	-0,47	-0,07
	Cit_PA	-1,80	-0,22	0,48	-16,81	-1,59	-4,72	-1,84	-20,54	-10,52
	PO4_PA	8,20	0,98	-2,18	76,34	7,22	21,43	8,34	93,27	47,77
	Acet_PA	-2,09	-0,25	0,56	-19,45	-1,84	-5,46	-2,12	-23,76	-12,17
	CaCO3_PA	-0,38	-0,05	0,10	-3,58	-0,34	-1,00	-0,39	-4,37	-2,24
	PERM1_SOD	-0,02	0,00	0,00	-0,16	-0,02	-0,04	-0,02	-0,19	-0,10
Group 3	F1	0,30	0,40	0,16	0,13	-0,03	0,17	0,41	0,19	0,05
	F2	0,06	0,06	-0,12	0,02	0,02	0,03	0,15	-0,18	0,38
	Mg_Ca_pc	0,06	0,01	0,01	0,79	-0,01	0,18	0,07	0,11	0,65
	Cl_TO	4,00	0,41	-1,05	46,02	1,11	10,30	5,15	-14,62	170,04
	PO4_TO	-2,85	-0,42	-6,40	-47,85	5,96	-10,45	0,17	-73,63	383,50
	K_pc_TO	4,34	0,47	0,63	53,63	-0,46	11,93	4,63	4,98	59,30
	Hg_TO	0,21	0,02	0,09	2,69	-0,08	0,60	0,19	0,97	-1,60
	Pb_TO	0,80	0,09	0,49	10,71	-0,44	2,37	0,65	5,34	-15,58
	Cd_TO	1,43	0,16	0,26	17,83	-0,20	3,97	1,50	2,28	15,76
	NO3_TO.1	2,70	0,31	1,00	34,67	-0,86	7,69	2,56	10,21	-5,76
Group 4	F1	-0,67	-1,02	0,10	-0,12	0,11	-0,09	-0,74	-0,95	-0,46
	F2	0,17	0,10	0,24	0,22	0,10	0,18	0,23	0,25	-0,34
	FDENSC	0,16	0,02	0,00	1,34	-0,04	0,19	0,13	1,06	1,01
	HC_BIN	-0,01	0,01	-0,09	-3,26	-0,20	-0,47	-0,03	-0,11	4,16
	PDEF	-0,07	-0,01	0,01	-0,36	0,03	-0,05	-0,06	-0,48	-0,80

Equation 2

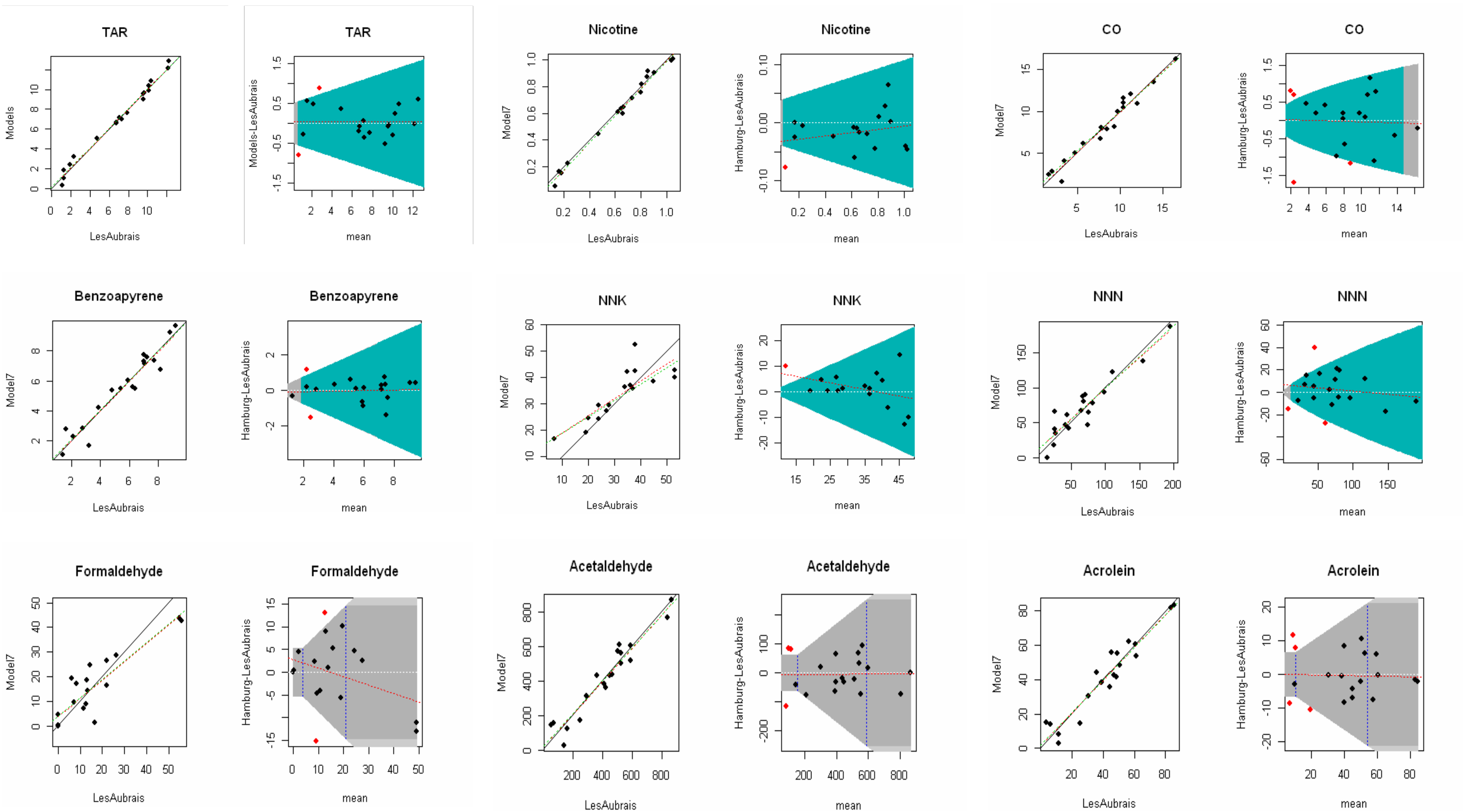
		NFDPM	Nicotine	CO	Acetaldehyde	Acrolein	Formaldehyde	BaP	NNK	NNN
Group 7	F1	-0,13	-0,13	-0,08	-0,11	-0,10	-0,04	-0,22	-0,38	0,13
	F2	-0,12	-0,20	0,01	0,02	0,02	0,17	-0,07	-0,37	-0,48
	F3	0,06	0,22	-0,15	0,06	0,13	0,18	0,12	0,14	-0,60
	TAR	0,05	0,01	0,01	2,17	0,24	0,07	0,05	0,55	-0,77
	NICO	0,78	0,13	-0,47	32,32	4,77	2,55	0,79	8,61	-25,48
	CO	0,00	-0,01	0,12	0,87	-0,16	-0,24	0,00	0,02	2,37
	Acetal_MS	0,00	0,00	0,00	0,03	0,00	0,00	0,00	0,01	0,04
	Acro_MS	0,00	0,00	0,02	0,32	-0,01	-0,03	0,00	0,04	0,32
	Fo_MS	0,00	0,00	0,00	0,46	0,05	0,05	0,01	0,02	-0,44
	BaP_MS	0,07	0,01	-0,03	3,70	0,50	0,32	0,08	0,73	-3,05
	NNK_MS	0,01	0,00	0,00	0,05	0,00	-0,07	0,01	0,16	0,52
	NNN_MS	0,00	0,00	0,00	-0,07	-0,01	-0,03	0,00	0,04	0,24
Group 5	F1	0,50	0,43	0,60	0,50	0,51	0,51	0,33	-0,04	0,61
	F2	-0,01	-0,05	-0,04	0,08	0,08	0,25	0,00	0,01	-0,57
	FV	-0,06	0,00	-0,07	-3,45	-0,36	-0,25	-0,03	0,02	-0,92
	PD	0,05	0,00	0,05	4,65	0,49	0,58	0,02	0,00	-1,45
	PDFNE	-0,09	-0,01	-0,11	-4,60	-0,48	-0,26	-0,04	0,03	-2,02

Coefficients of exogenous variables in Hoffmann compounds models (from model 7)

How to assess prediction quality of Hoffmann Compounds?

Application to cigarette data

Hoffmann compounds: **1) laboratory measure *vs* model 7 prediction; 2) Relative error / reproducibility limits**



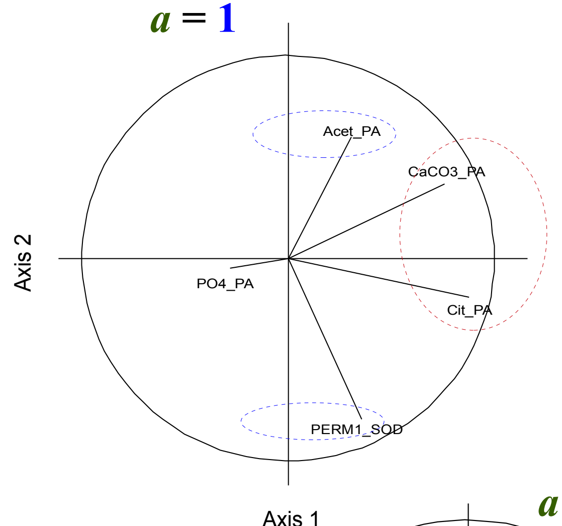
Application to cigarette data

Multiple Costructure criterion: effect of exponent *a*

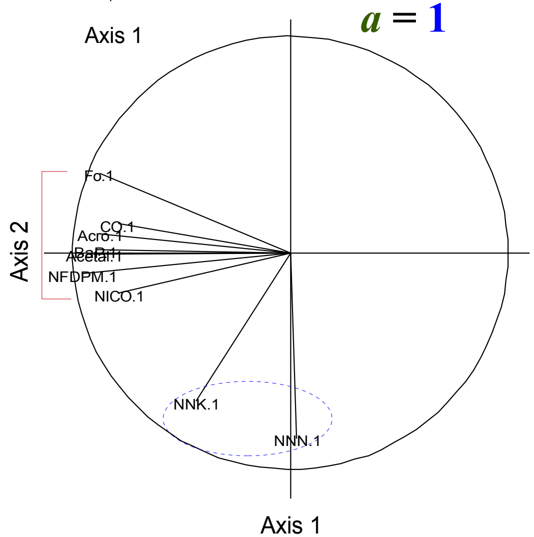
***Model* = 2 2 2 2 2 3 3** $\quad$ ***a* = 1, … , 7**

Groups 1, 3, 4, 5, 6 → Very little change: **Important bundle structures are close to components**

Group 2: ***a* = 1**



Group 7: ***a* = 1**



Application to cigarette data

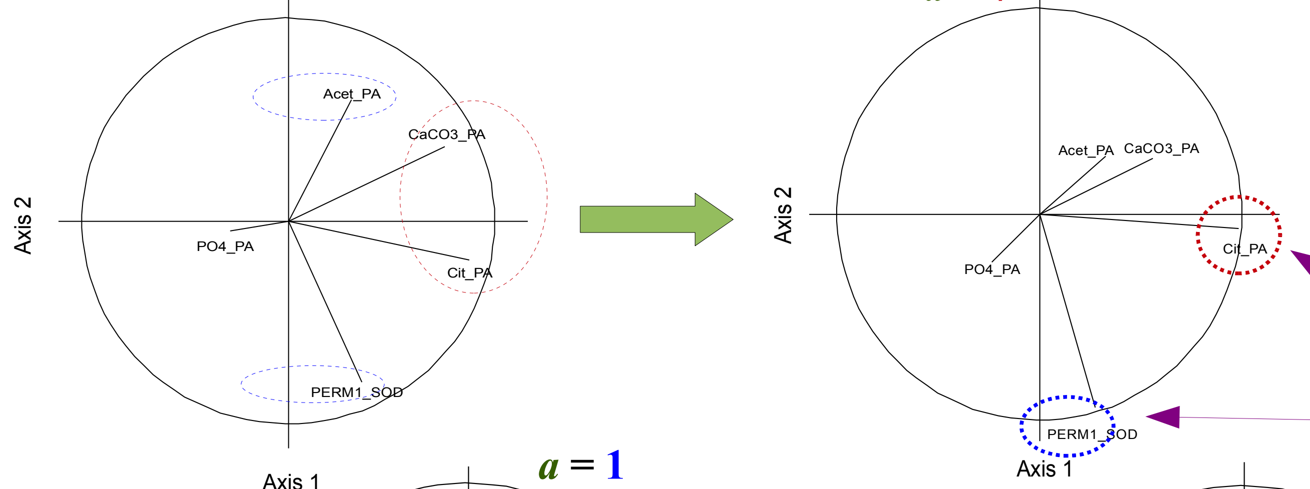
Multiple Costructure criterion: effect of exponent *a*

$Model = 2\ 2\ 2\ 2\ 2\ 3\ 3$ $\qquad a = 1, \dots, 7$

Groups 1, 3, 4, 5, 6 → Very little change: **Important bundle structures are close to components**

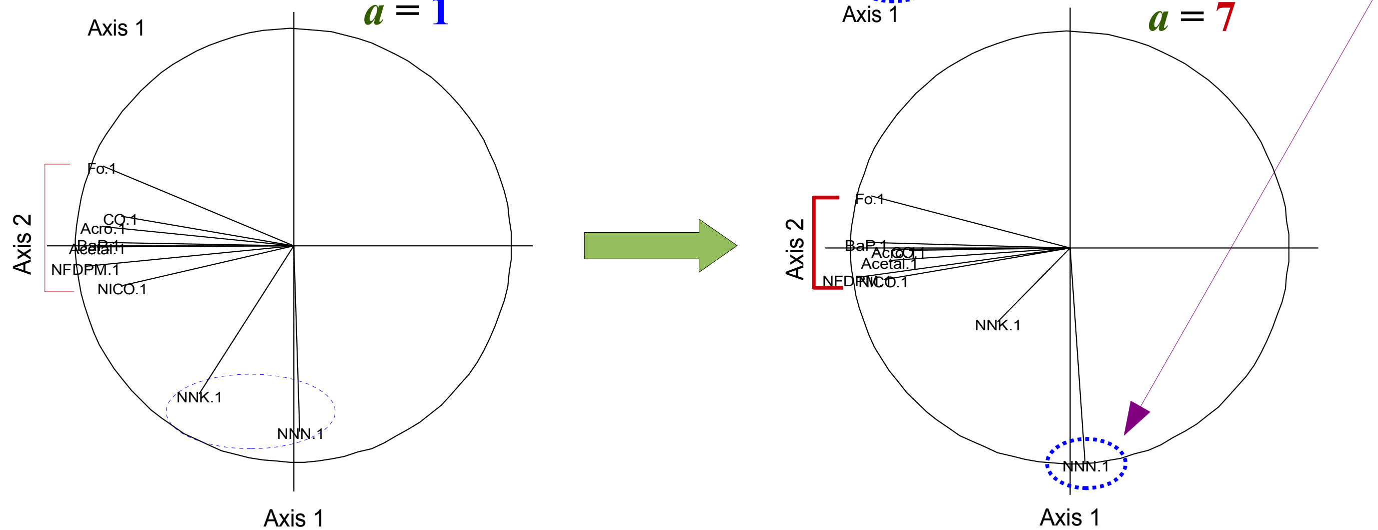
Group 2:

$a = 1$ $\qquad a = 7$



Group 7:

$a = 1$ $\qquad a = 7$



towards variable selection

Application to cigarette data

Conclusion

- From the ***explanatory*** point of view,

  THEME allowed to separate the **complementary roles**, on Hoffmann Compounds, of:

  - Tobacco quality (stalk position, pct of cutters and strips...)
  - Tobacco type (Burley, Flue Cured, Oriental, Virginia)
  - Combustion chemical enhancers or inhibitors related to tobacco or paper
  - Filter retention power.
  - Filter ventilation power

  When all predictors are mixed up, the filter ventilation effect masks the role of chemical constituents.

  THEME confirmed the relevance of the chemists' conceptual model.

- From the ***predictive*** point of view,

  THEME gave out a complete and robust model having accuracy within reproducibility limits

Application to cigarette data

Conclusion

- From the ***explanatory*** point of view,

  THEME allowed to separate the **complementary roles**, on Hoffmann Compounds, of:

  - Tobacco quality (stalk position, pct of cutters and strips...)
  - Tobacco type (Burley, Flue Cured, Oriental, Virginia)
  - Combustion chemical enhancers or inhibitors related to tobacco or paper
  - Filter retention power.
  - Filter ventilation power

  When all predictors are mixed up, the filter ventilation effect masks the role of chemical constituents.

  THEME confirmed the relevance of the chemists' conceptual model.

- From the ***predictive*** point of view,

  THEME gave out a complete and robust model having accuracy within reproducibility limits

Software

Free **R-based** **User-friendly interface**

Beta THEME 1.0 available on (mail) demand

Thank you, all