

# Regularization approaches in sensory descriptive analysis

Applications to Kokumi-product development

Eduard Derks, Kae Morita, Tetsuo Aishima

Sensometrics  
Juli 12th, 2012

# NPD-study コク味

- Goal of project
  1. Understand Japanese concept
  2. Identify koku-drivers in savoury/cullinary applications
  3. Support NPD to developed a koku-enhancing compound
- Sub-goal (method development)

Develop strategies to

  - link koku to sensory data (difference tests, QDA)
  - deal with large datasets
  - improve interpretability

# Kokumi

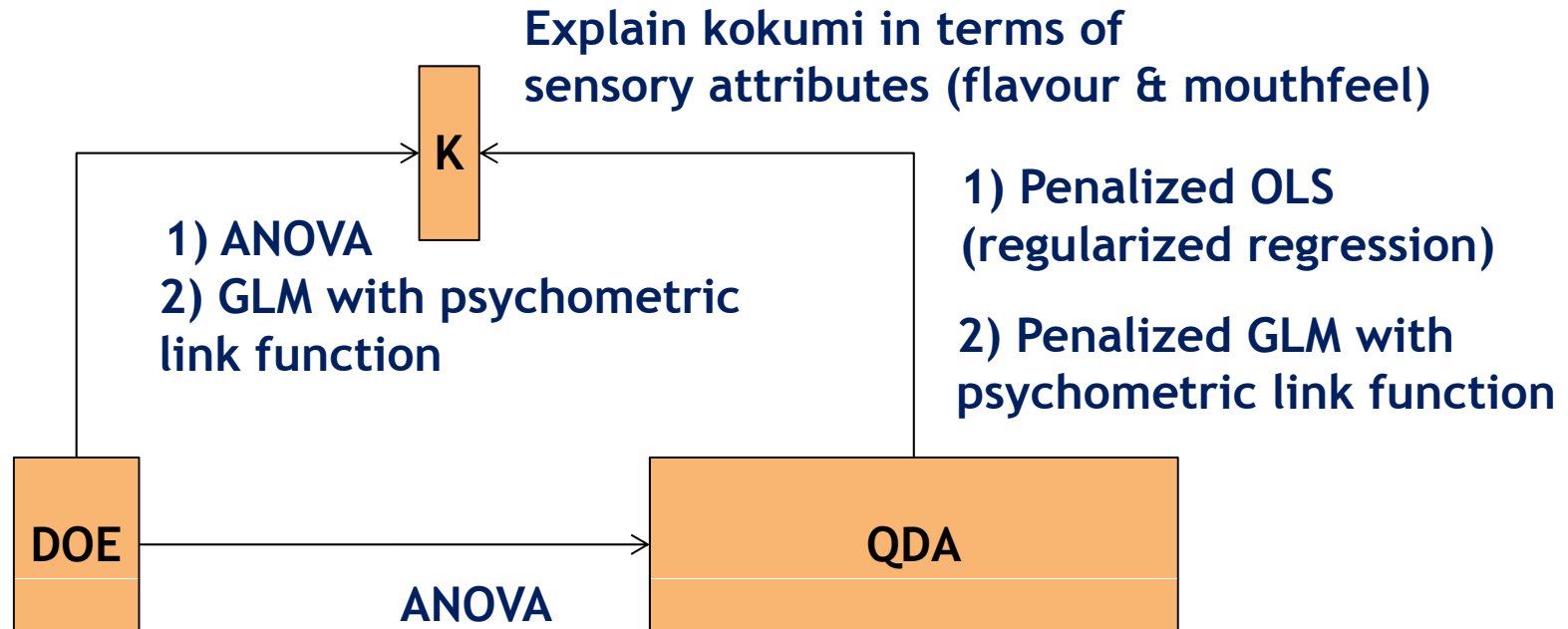
Koku = コク = 濃香 = thickness, complexity, depth and/or width  
mi = 味 = 味道 = taste



# Project steps

- Evaluation in expert teams and QDA-panels (Japanese & Dutch)
- Interviews with Japanese cooks (n=80)
- Interviews with APAC consumers (Japan, China, Korea, n=600)
- Meat stew sensomics (NMR, GPC, LCMS, GCMS, QDA, etc.)
- YE-based reconstruction of koku-targeted flavour & mouthfeel
- Consumer tests: testing NPD samples against competition
- Lots of data - difficult to interpret - need for sparse methods both for regression and explorative analysis

# Data & models



## Alternatives

### 1) PLS/RR

- greedy
- developed for smooth spectra

### 2) GLM with psycho-lf. on PCA scores

- more difficult to interpret

# Regularization in multivariate regression

- Multivariate regression

$$y = \mathbf{X}\beta + \varepsilon$$

- OLS minimizes

$$\|y - \mathbf{X}\beta\|^2$$

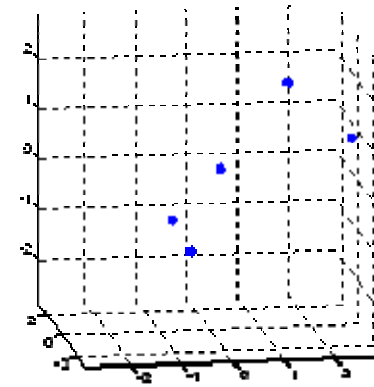
- Solution

$$\beta = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t y$$

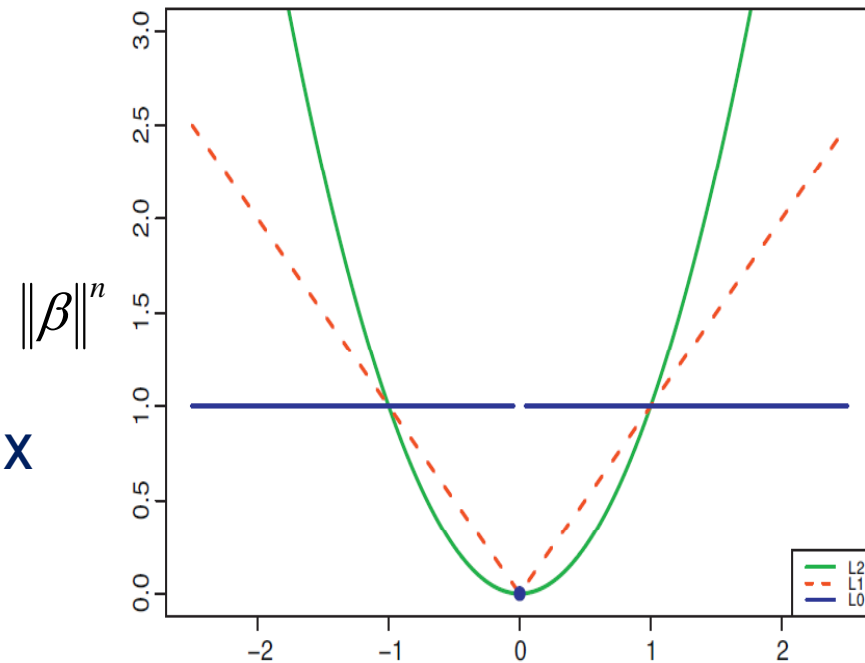
- Inversion problems for collinear X

- Solution: restrict beta's

$$\|y - \mathbf{X}\beta\|^2 + \lambda \|\beta\|^n$$



L2, L1 and L0 penalty function



# Regularization

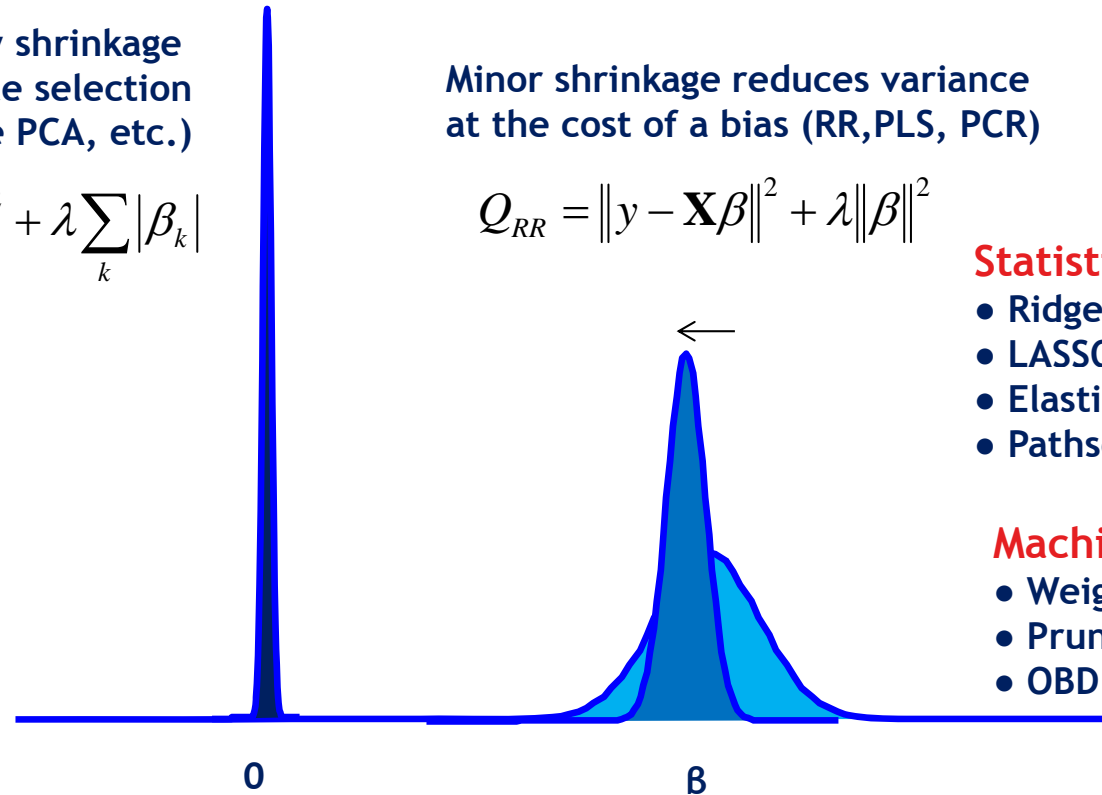
## in multivariate regression

Heavy shrinkage  
brings variable selection  
(Sparse OLS, Sparse PCA, etc.)

$$Q_{L1} = \|y - \mathbf{X}\beta\|^2 + \lambda \sum_k |\beta_k|$$

Minor shrinkage reduces variance  
at the cost of a bias (RR, PLS, PCR)

$$Q_{RR} = \|y - \mathbf{X}\beta\|^2 + \lambda \|\beta\|^2$$



### Statistics

- Ridge regression
- LASSO/ LARS
- Elastic net
- Pathseeker

### Machine learning

- Weight decay
- Pruning
- OBD

# Regularization

## in principal component analysis

- PCA

$$\hat{\mathbf{X}} = tp^t$$

$$Q_{PCA} = \|\mathbf{X} - tp^t\|^2$$

- Issues:

messy biplots

many variables (e.g. internal prefmapping, higher order QDA)

many objects (e.g. consumer studies)

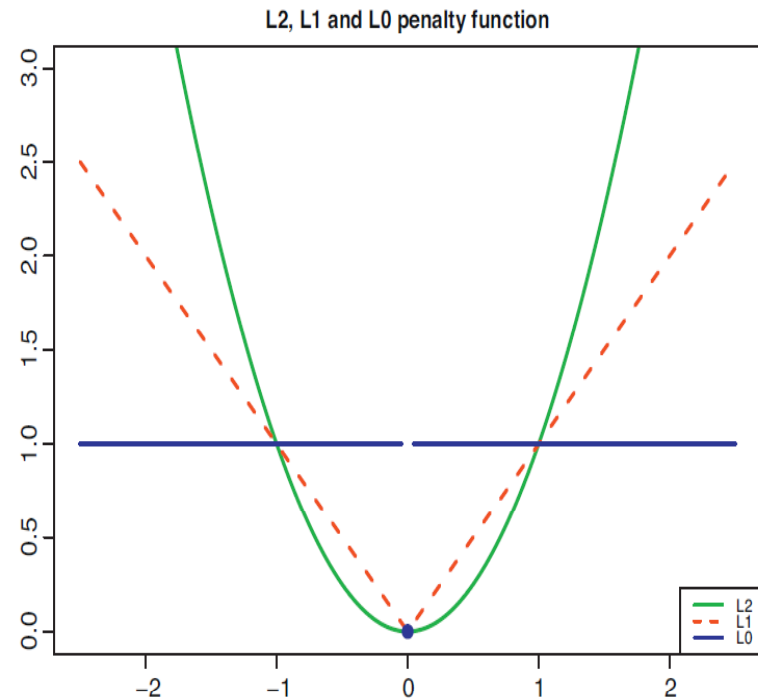
- Sparse PCA

Shrink loadings (automatic attribute selection)

$$Q_{L2(p)} = \|\mathbf{X} - tp^t\|^2 + \lambda \|p\|^2$$

$$Q_{L1(p)} = \|\mathbf{X} - tp^t\|^2 + \lambda \sum |p_j|$$

$$Q_{L0(p)} = \|\mathbf{X} - tp^t\|^2 + \lambda \sum_j \mathbb{1}(|p_j| > 0)$$

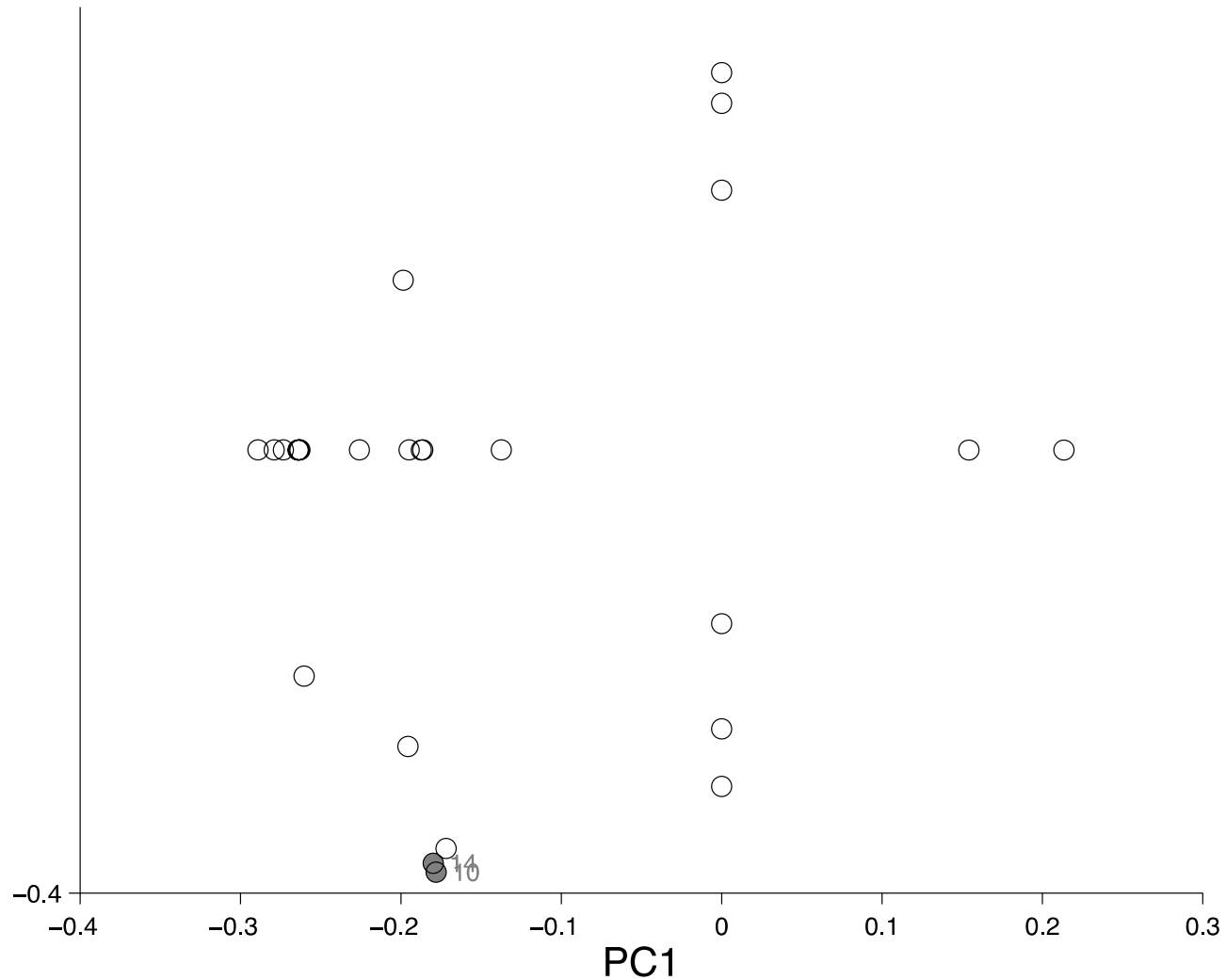




# Regularization

in principal component analysis

- Easier to interpret
- due to sparseness
  - due to the fact that PC's get a meaning



# Regularization

## in principal component analysis

- Can also be applied for sparse scores
- or penalization on scores and loadings simultaneously

This is called CoClustering\*

$$Q_{coclustering} = \|\mathbf{X} - tp^t\|^2 + \lambda \sum_i |t_i| + \mu \sum_j |p_j|$$

- method of choice when specific variables are related by specific groups of objects
  - contingency (COUNT) tables
  - CATA biplots (consumers x checked attributes)
  - sparse scoring data

\* Bro, R., Papalexakis, E. E., Acar, E. and Sidiropoulos, N. D. (2012), *Coclustering—a useful tool for chemometrics*. *J. Chemometrics*, 26: 256–263.

	Has eyes	Number of	Carnivor	Feather	Wings	Domesticiz	Eaten by C	>100kg	>2m	Breathe ur	Extinct	Dangerous	Life expect
Giraffe	1	4	0	0	0	0	0	0	1	1	0	0	30
Cow	1	4	0	0	0	1	1	1	1	0	0	0	15
Lion	1	4	1	0	0	0	0	1	0	0	0	1	15
Gorilla	1	4	0	0	0	0	0	1	0	0	0	1	30
Fly	1										0	0	0.1
Spider	1										0	0	1
Shark	1										0	1	50
House	0										0	0	100
Horse	1										0	0	15
Elephant	1										0	0	35
Mammoth	1										1	0	35
Sabre Tige	1										1	1	15
Pig	1										0	0	25
Cod	1										0	0	40
Eel	1										0	0	55
Jellyfish	1										0	0	0.7
Dolphin	1										0	0	30
Nemo	1										0	0	1
Shrimp	1										0	0	1
Dog	1										0	0	13
Cat	1										0	0	25
Fox	1										0	0	14
Wolf	1										0	1	18
Rabbit	1										0	0	9
Chicken	1										0	0	15
Eagle	1										0	0	55
Seagull	1										0	0	10
Blackbird	1	2	1	1	1	0	0	0	0	0	0	0	18
Bat	1	2	1	0	1	0	0	0	0	0	0	0	24
T. Rex.	1	4	1	0	0	0	0	1	1	0	1	1	40
Neanderth	1	4	1	0	0	0	0	0	0	0	1	0	50
Triceratops	1	4	1	0	0	0	0	1	1	0	1	1	30

**Cluster 1**

Penguin

Blackbird Walk on two legs

Seagull Has a beak

Eagle Wings

Chicken Feather

**Cluster 2**

Triceratops

Neanderthal

T. Rex.

Sabre Tiger

Mammoth Extinct

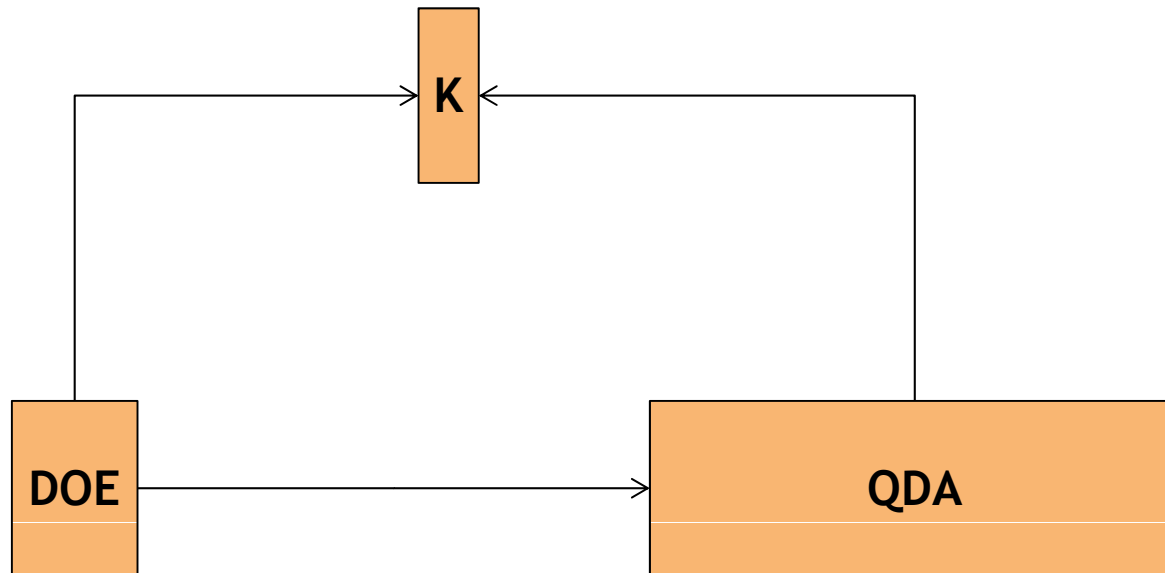
**Cluster 4**

Chicken

**Cluster 5**

Shrimp

## Now some applications ...

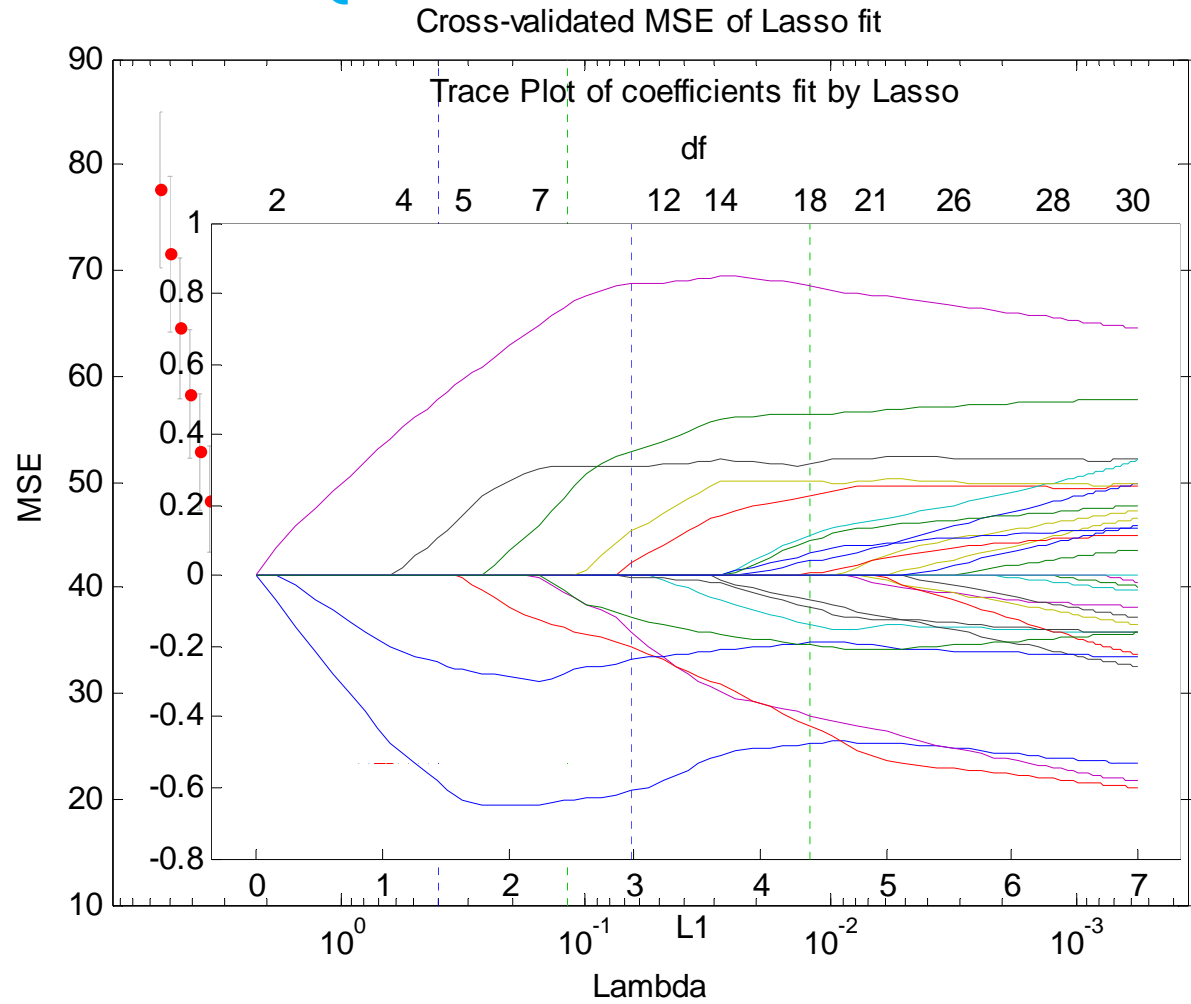


# LASSO

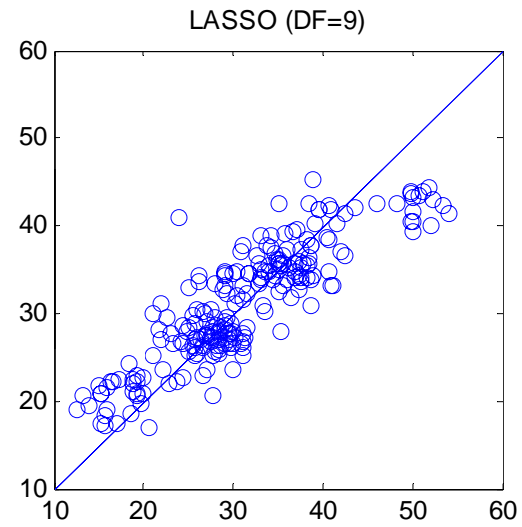
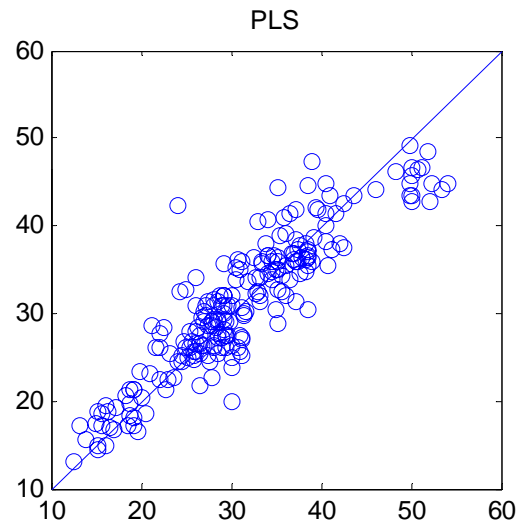
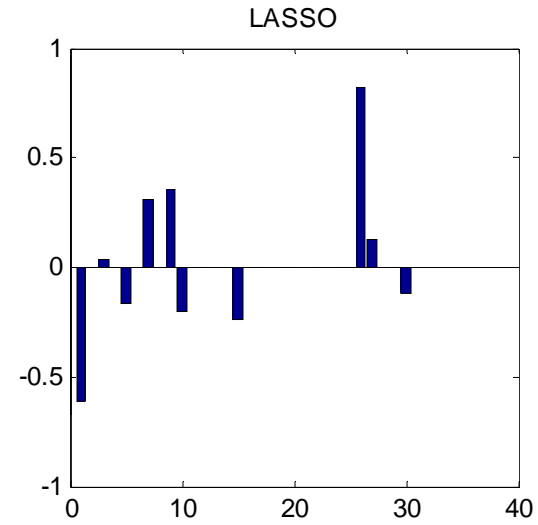
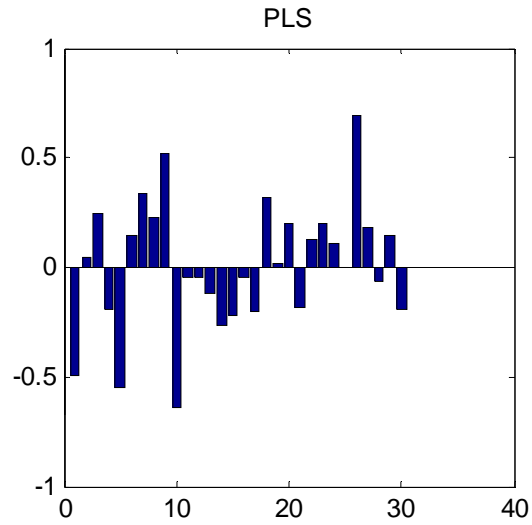
## koku versus QDA

K

QDA

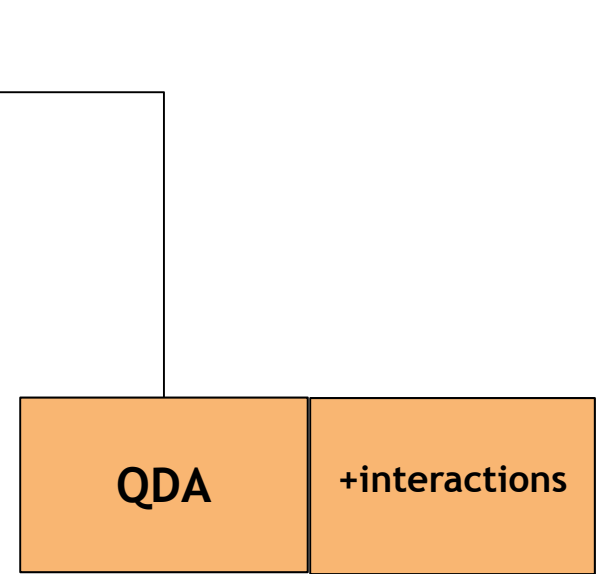
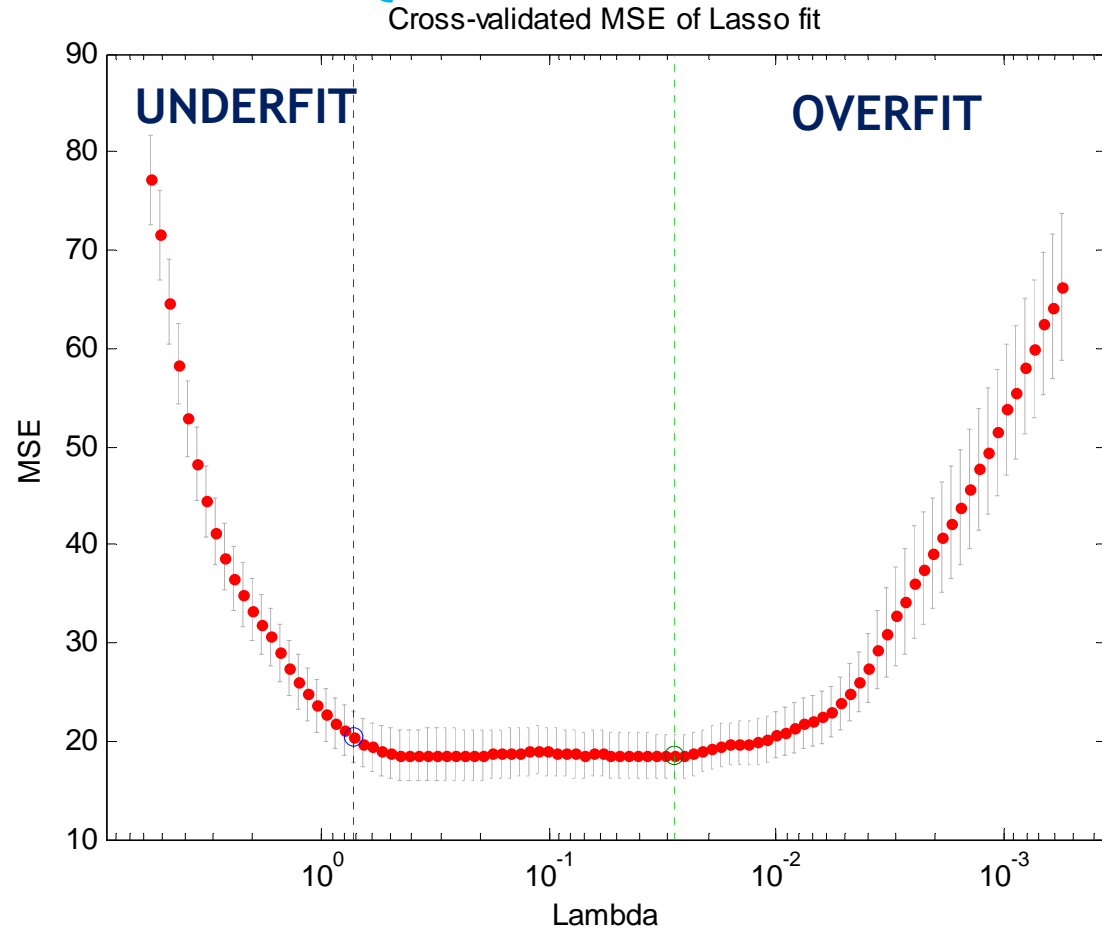


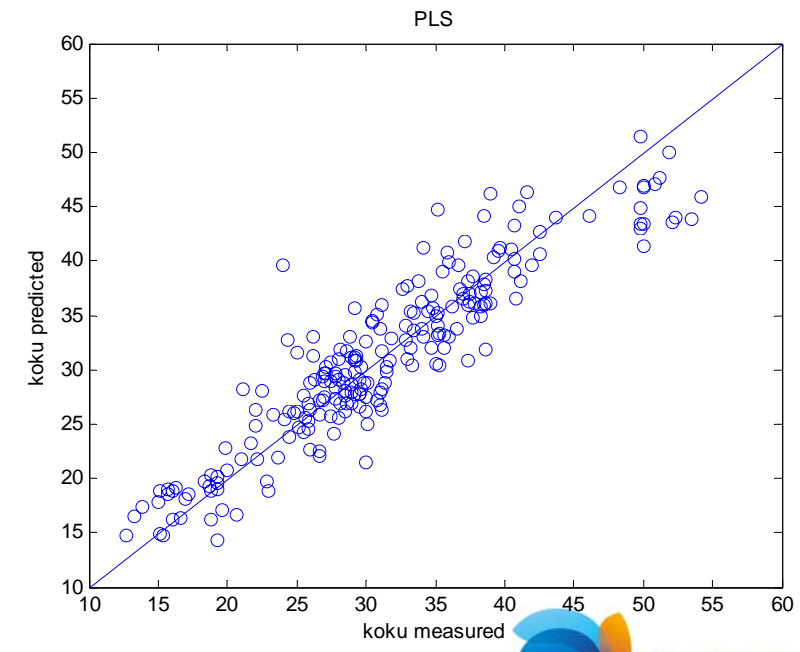
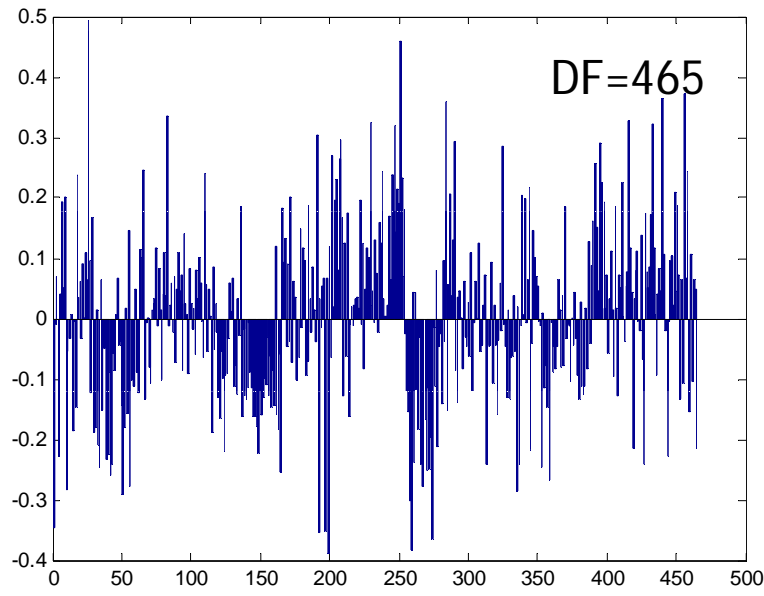
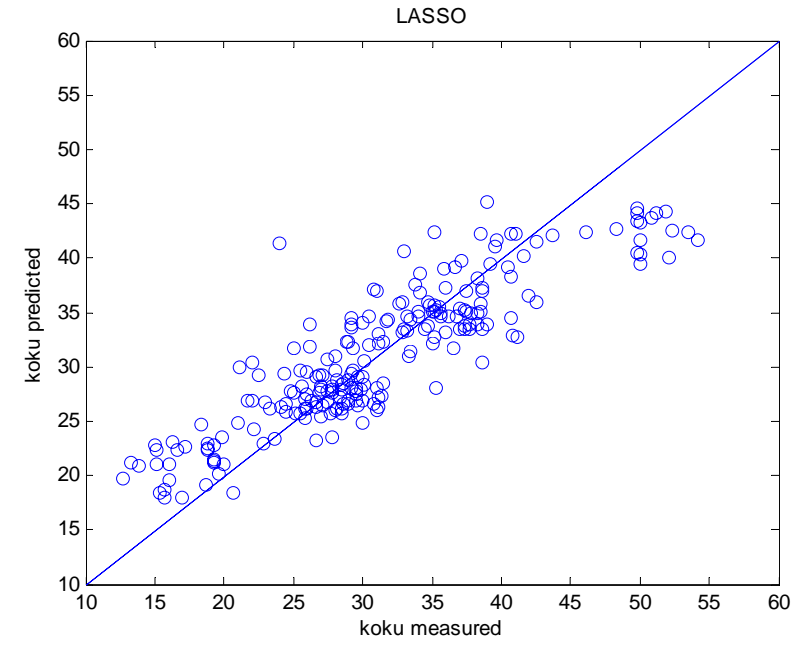
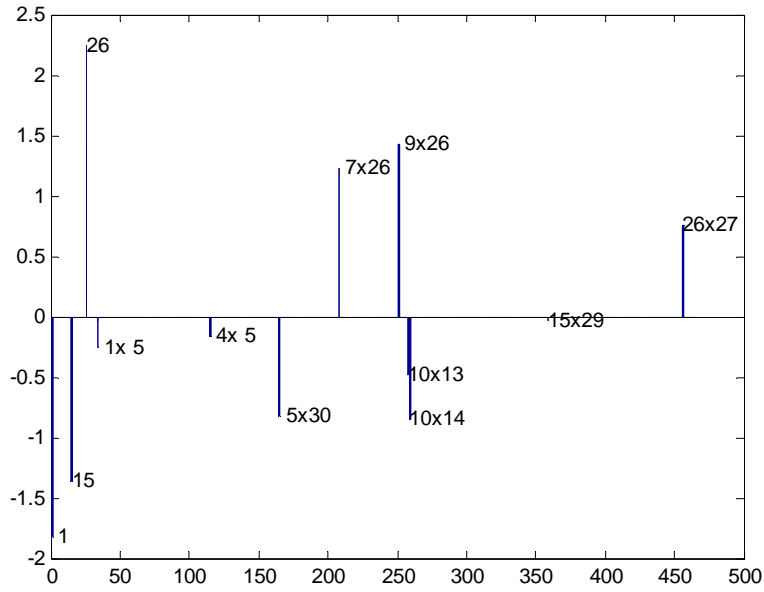
# LASSO



# LASSO

## koku versus QDA



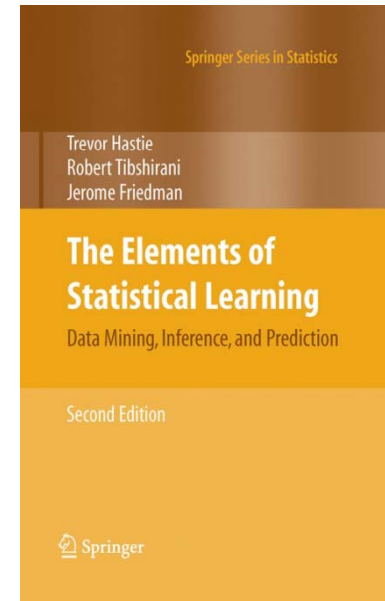




# First conclusions

- LASSO selects a limited set of sensory attributes and interaction terms to explain kokumi with similar predictability compared to PLS
- But, LASSO improves interpretation (variable selection)
- PLS overfits (not demonstrated here)
- PLS (and RR) do not handle mega/multivariate sensory data well
  - eager to overfit
  - seeks correlated coefficients\*

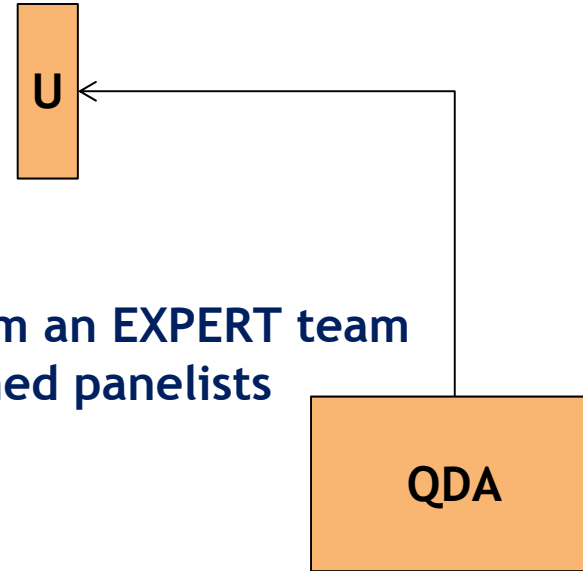
*\* Hastie, T., Tibshirani, R., Friedman, J. 'The elements of statistical learning'. Springer series in Statistics. Page 83*



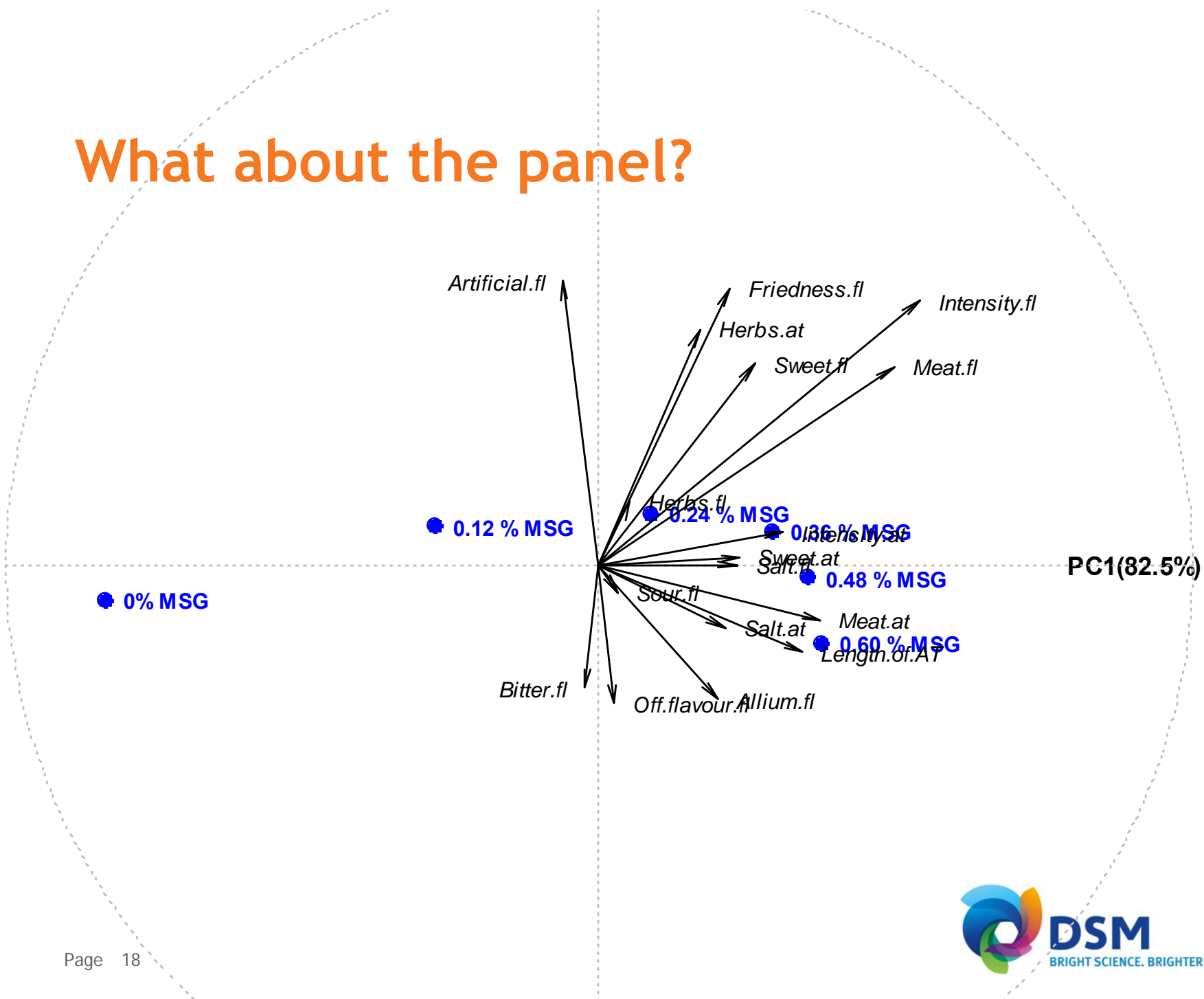
# LASSO-GLM

## MSG dose-response analysis

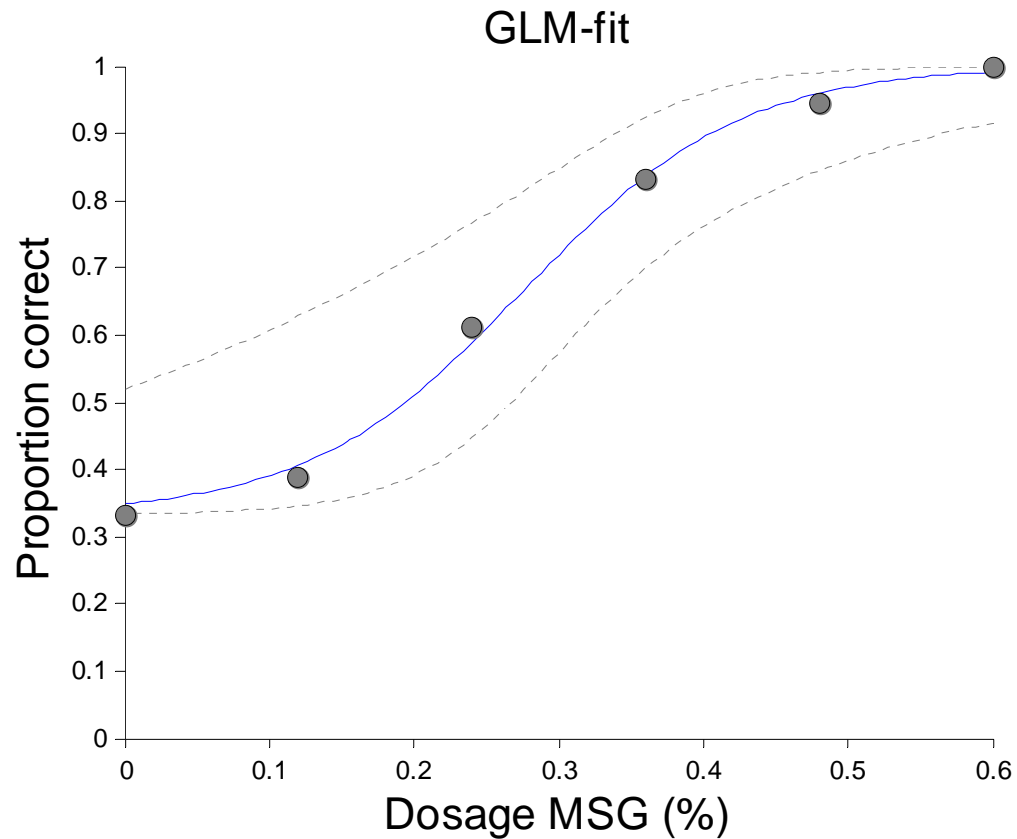
- Purpose: model umami perception from an EXPERT team as a function of QDA attributes with trained panelists
- EXPERT team (n=6) used 3AFC (2reps)
- QDA panel (n=14) (no umami in attribute list)
- MSG series (0, 0.12, 0.24, 0.36, 0.48, 0.60 percent)
- Applied in beef bouillon
- Proportion correct odds were modeled with GLM as function of MSG dosage and QDA attributes
- GLM model was organized with a 3AFC-psychometric link function and L1 penalty



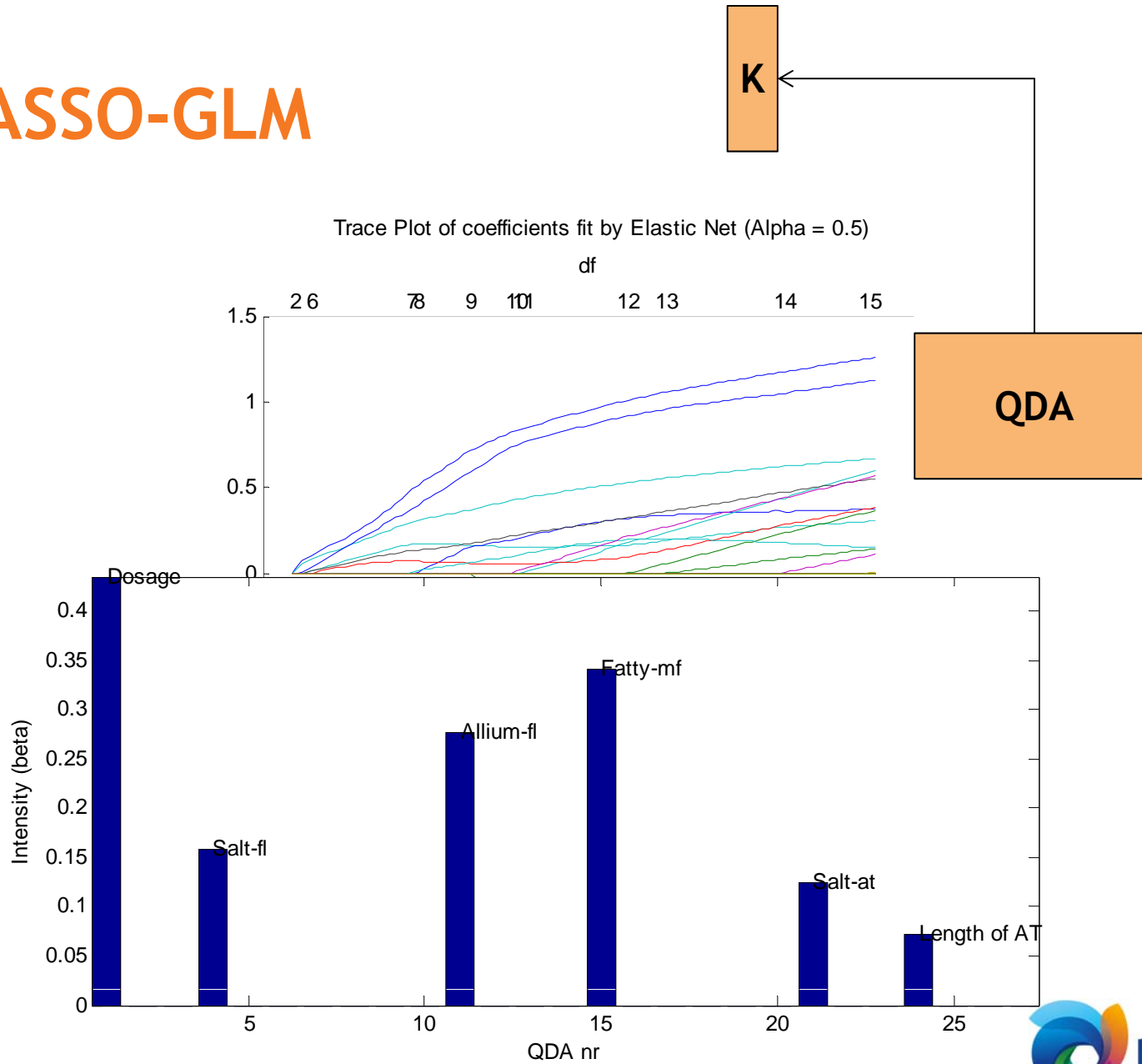
# What about the panel?



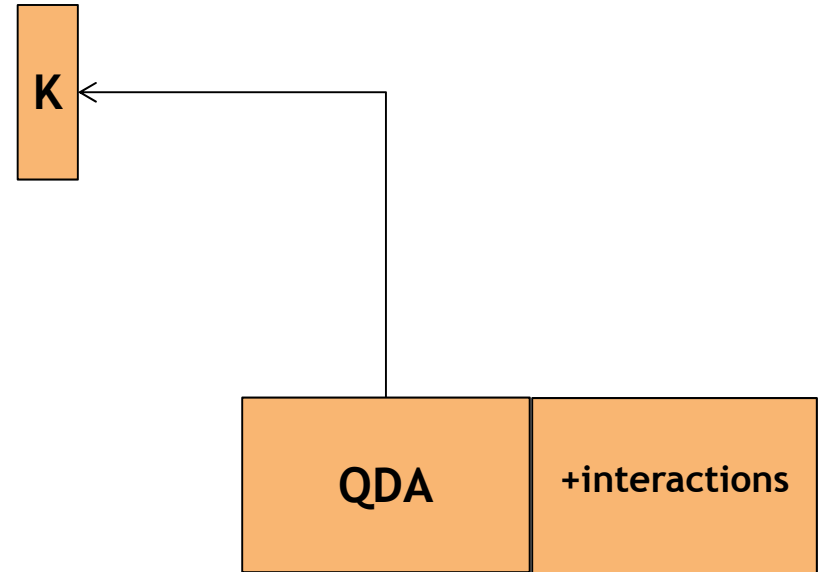
# What about the experts?



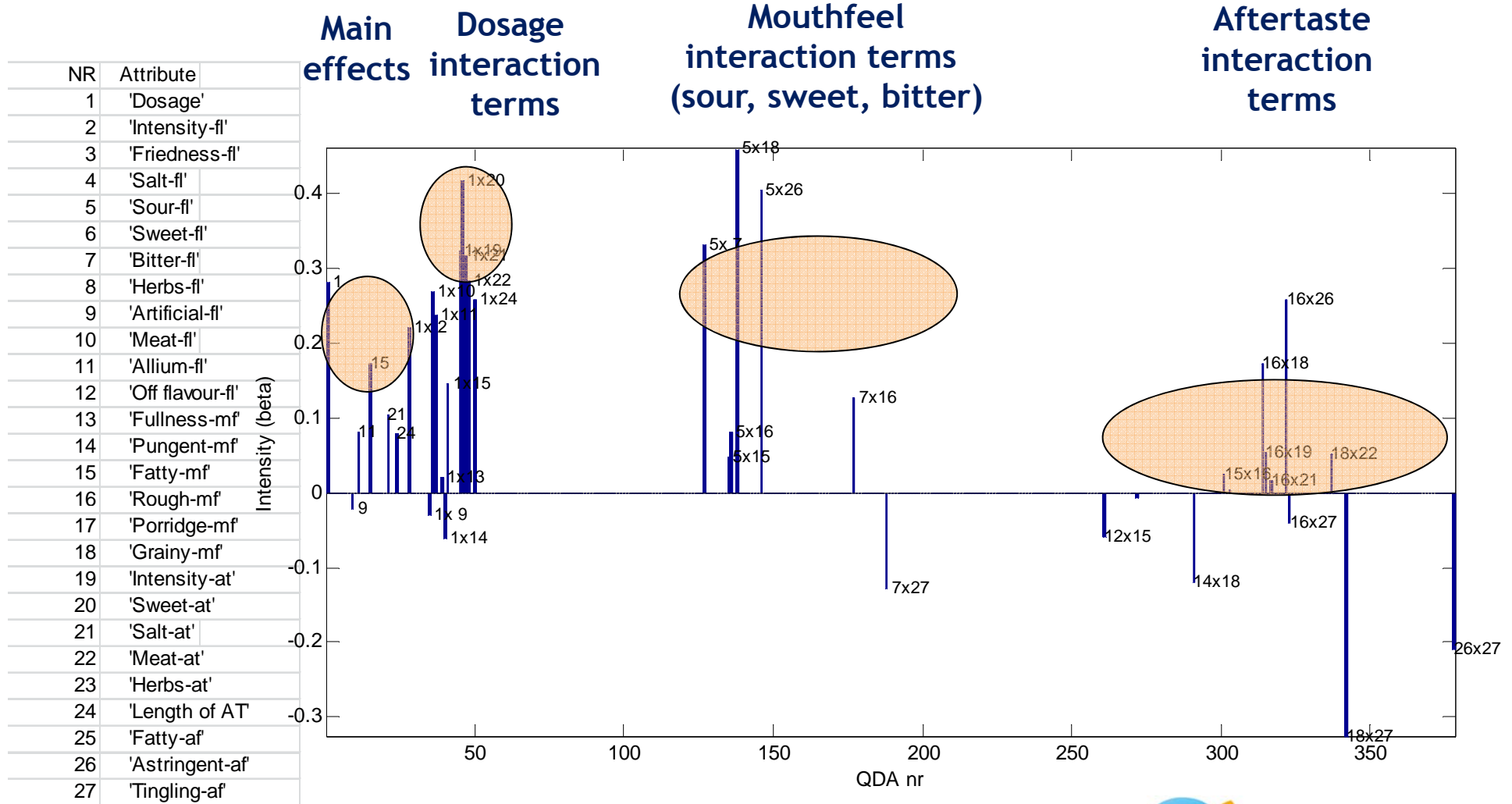
# LASSO-GLM



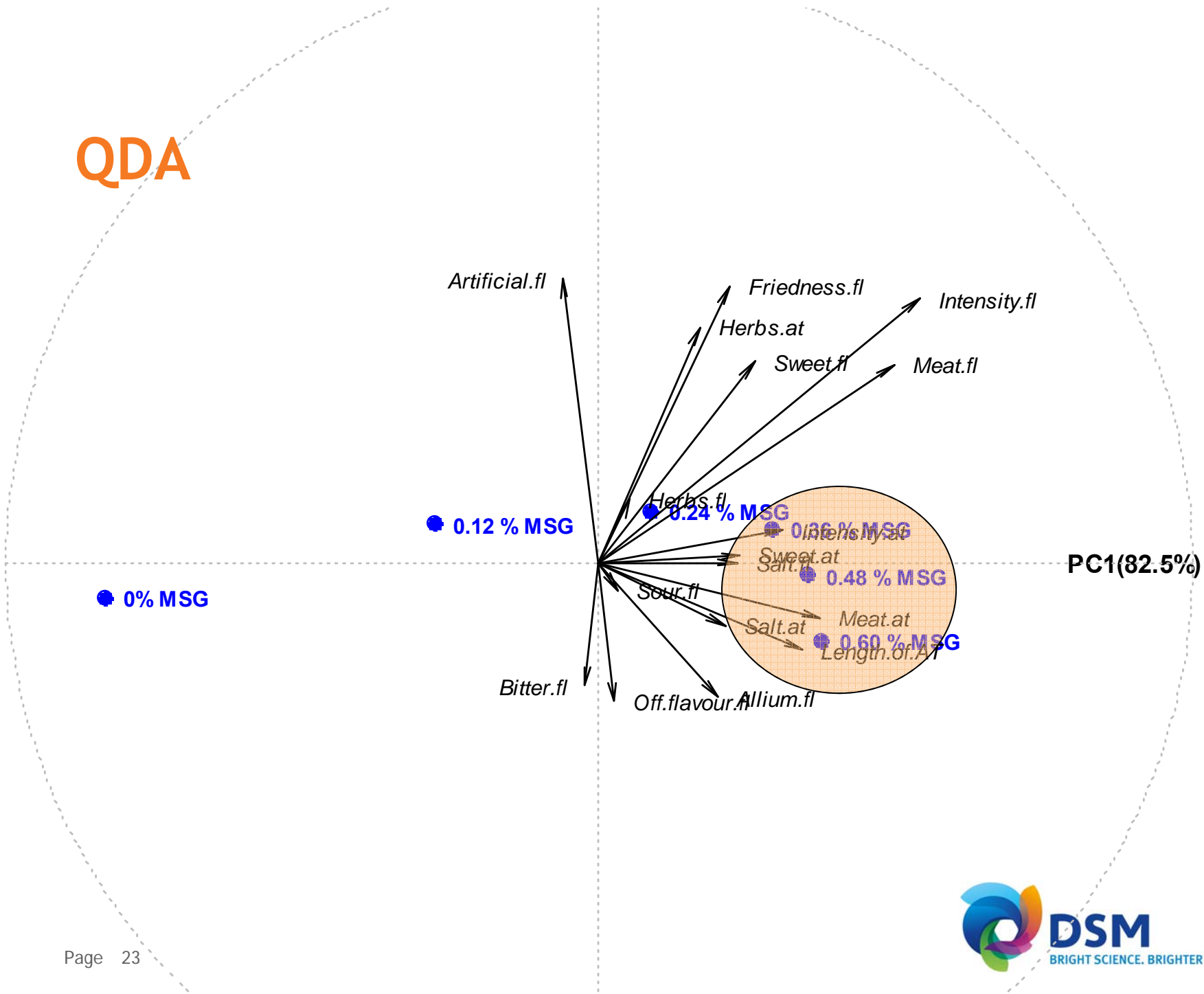
# LASSO-GLM



# LASSO-GLM



# QDA





# Take home messages / conclusion

- Regularization is a powerful concept that enables
  - stabilization (RR)
  - sparseness (variable selection in PCA in OLS)
  - improved interpretability
  - improved predictability (for unseen data)
- Other types of restrictions (i.e. other than  $L_n$ ) can be implemented as well
  - roughness penalties for smoothness (Smooth PCA)
  - dispersion penalties (clustering)
  - m-univariate regression (using correlation in y) (applications in shelf life analysis)

Thank you for your attention!

**BRIGHT SCIENCE. BRIGHTER LIVING.™**