# Applications of Response Dimension Reduction in Large $p$-Small $n$ Problems

Minjee Kim[1]\*, Jae Keun Yoo[1]

**[1] Department of Statistics
Ewha Womans University, Korea**

Korean Data Analysis Society, December 6, 2024

# Response Dimension Reduction

- **Sufficient Dimension Reduction**
  - Replace $\mathbf{X} \in \mathbb{R}^p$ with lower dimensional projection $\mathbf{P}_\mathcal{S}\mathbf{X}$, without losing information of $\mathbf{Y} \mid \mathbf{X}$
  - Find the orthogonal projection operator $\mathbf{P}_\mathcal{S}$, such that

    $$\mathbf{Y} \perp\!\!\!\perp \mathbf{X} \mid \mathbf{P}_\mathcal{S}\mathbf{X}$$

- **Response Dimension Reduction**
  - In multivariate regression of $\mathbf{Y} \in \mathbb{R}^r \mid \mathbf{X} \in \mathbb{R}^p$ ($r \geq 2$), **Sample size** $n$ could be smaller than $p$ or $r \Rightarrow$ **Large *p*-Small *n* Problem**
    - Lead to exponential increase in the number of parameters (**curse of dimensionality**)
  - Reduce the dimension of $\mathbf{Y}$ without loss of information on $E(\mathbf{Y} \mid \mathbf{X})$

# Non-parametric Method (YC-method)

## Linear Response Reduction (Yoo and Cook, 2008)

$$E(\mathbf{Y}|\mathbf{X}) = E\{\mathbf{P}_{\mathsf{L}(\Sigma_y)}^{\mathsf{T}} \mathbf{Y}|\mathbf{X}\}$$

where $\mathbf{L} \in \mathbb{R}^{r \times q}$, $q \leq r$, and $\mathbf{P}_{\mathsf{L}(\Sigma_y)} = \mathbf{L}(\mathbf{L}^{\mathsf{T}}\Sigma_y\mathbf{L})^{-1}\mathbf{L}^{\mathsf{T}}\Sigma_y$ is an **orthogonal projection operator**.

## Conditional Response Reduction (Yoo and Cook, 2008)

$$E(\mathbf{Y}|\mathbf{X}) = E\{E(\mathbf{Y}|\mathbf{X}, \mathbf{K}^{\mathsf{T}}\mathbf{Y})|\mathbf{X}\} = E\{E(\mathbf{Y}|\mathbf{K}^{\mathsf{T}}\mathbf{Y})|\mathbf{X}\} = E\{g(\mathbf{K}^{\mathsf{T}}\mathbf{Y})|\mathbf{X}\}$$

where $\mathbf{K} \in \mathbb{R}^{r \times k}$, $k \leq r$, $\mathbf{K} \neq \mathbf{I}_r$, and $g(\cdot)$ represents an **unknown function**.

- Column spaces of **L**, **K** : *Response dimension reduction subspace*
- $\mathcal{S}(\mathbf{K}) \subseteq \mathcal{S}(\mathbf{L})$, and $\mathcal{S}(\mathbf{K}) = \mathcal{S}(\mathbf{L})$ iff $E(\mathbf{Y}|\mathbf{K}^{\mathsf{T}}\mathbf{Y} = a)$ is linear in $a$
- To estimate **L** and **K**, $\Sigma_y^{-1}\mathrm{cov}(\mathbf{Y}, \mathbf{X}^{\mathsf{T}})\Sigma_x^{-1}$ is needed
  - $\Sigma_x = \mathrm{cov}(\mathbf{X})$ and $\Sigma_y = \mathrm{cov}(\mathbf{Y})$ are positive-definite matrices

# Principal Response Reduction

## Principal Response Reduction (Yoo, 2018)

Assume both $E(\mathbf{Y})$ and $E(\mathbf{X})$ are zero without loss of generality:

$$\mathbf{Y} = \mathbf{\Gamma}\boldsymbol{\nu}_x + \varepsilon$$

where $d \leq r$, $\mathbf{\Gamma} \in \mathbb{R}^{r \times d}$ is an **orthogonal** matrix, $\varepsilon \sim N(0, \mathbf{\Sigma})$, and $\mathrm{cov}(\boldsymbol{\nu}_x, \varepsilon) = 0$.
$\boldsymbol{\nu}_x$ is a $d$-dimensional **unknown function** of $\mathbf{X}$, and $\mathbf{\Sigma}_x \boldsymbol{\nu}_x = 0$.

- $\mathcal{S}(\mathbf{\Gamma})$ is invariant, and reduces subspace for $\mathbf{\Sigma}$
- $\mathbf{\Sigma} = \mathbf{\Gamma}\mathbf{\Omega}\mathbf{\Gamma}^{\mathsf{T}} + \mathbf{\Gamma}_0\mathbf{\Omega}_0\mathbf{\Gamma}_0^{\mathsf{T}}$
  - $\mathbf{\Gamma}_0 \in \mathbb{R}^{r \times (r-d)}$, $\mathbf{\Gamma}_0^{\mathsf{T}}\mathbf{\Gamma}_0 = \mathbf{I}_{r-d}$, $\mathbf{\Gamma}_0^{\mathsf{T}}\mathbf{\Gamma} = 0$, $\mathbf{\Omega} = \mathbf{\Gamma}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{\Gamma}$, and $\mathbf{\Omega}_0 = \mathbf{\Gamma}_0^{\mathsf{T}}\mathbf{\Sigma}\mathbf{\Gamma}_0$
- $E(\mathbf{Y}|\mathbf{X}) = E(\mathbf{P}_{\mathbf{\Gamma}(\mathbf{\Sigma}_y)}^{\mathsf{T}}\mathbf{Y}|\mathbf{X})$ holds
- MLE for $\mathbf{\Gamma}$ is derived from the 1st $d$ largest eigenvectors of $\hat{\mathbf{\Sigma}}_y$
  - $\mathcal{L}_{\mathrm{PRR}}(\mathbf{\Gamma}_0, \mathbf{\Upsilon}) = -\frac{n}{2}\log|\mathbf{\Gamma}_0^{\mathsf{T}}\hat{\mathbf{\Sigma}}_y\mathbf{\Gamma}_0| - \frac{n}{2}\log|\mathbf{\Upsilon}|$, where $\mathbf{\Upsilon} \in \mathbb{S}^{r \times r}$ stand for $\mathbf{\Omega}$
  - $\Rightarrow$ No need for inverse matrix
- However, in **PRR**, $\mathbf{\Gamma}$ does not consider $\mathbf{X}$, only utilizes $\mathbf{Y}$

# Principal Fitted Response Reduction

### Principal Fitted Response Reduction (Yoo, 2018)

To incorporate **X**, assume $\nu_x = \psi \mathbf{f}_x$,

$$\mathbf{Y} = \mathbf{\Gamma}\psi\mathbf{f}_x + \varepsilon$$

where $\psi \in \mathbb{R}^{d \times q}$ is an unknown matrix, $\mathbf{f}_x$ is a $q$-dimension **known vector function** of **X**, and $\mathbf{\Sigma}_x\mathbf{f}_x = 0$.

- $\mathbb{Y}$: $n \times r$ response data matrix
- $\mathbb{F}$: $n \times q$ matrix stacking $\mathbf{f}_x^\intercal$
- $\mathbf{P}_{\mathbb{F}} = \mathbb{F}(\mathbb{F}^\intercal\mathbb{F})^{-1}\mathbb{F}^\intercal$
- $\hat{\mathbf{\Sigma}}_{\text{fit}} = \mathbb{Y}^\intercal\mathbf{P}_{\mathbb{F}}\mathbb{Y}/n$ and $\hat{\mathbf{\Sigma}}_{\text{res}} = \hat{\mathbf{\Sigma}}_y - \hat{\mathbf{\Sigma}}_{\text{fit}}$
- MLE for **Γ** does not have a closed-form solution
  - $\mathcal{L}_{\text{PFRR}}(\mathbf{\Gamma}, \mathbf{\Gamma}_0) = -\frac{n}{2}\log|\mathbf{\Gamma}_0^\intercal\hat{\mathbf{\Sigma}}_y\mathbf{\Gamma}_0| - \frac{n}{2}\log|\mathbf{\Gamma}^\intercal\hat{\mathbf{\Sigma}}_{\text{res}}\mathbf{\Gamma}|$
  - **Γ** is influenced by both $\hat{\mathbf{\Sigma}}_y$ and $\hat{\mathbf{\Sigma}}_{\text{res}}$
  - $\Rightarrow$ Only needs inverse of $\mathbb{F}^\intercal\mathbb{F}$

# Unstructured Principal Fitted Response Reduction

## Unstructured Principal Fitted Response Reduction (Yoo, 2019)

Assume that $\varepsilon \sim N(0, \mathbf{\Sigma} > 0)$, and $\operatorname{cov}(\boldsymbol{\nu}_x, \varepsilon) = 0$:

$$\mathbf{Y} = \mathbf{\Gamma}\boldsymbol{\psi}\mathbf{f}_x + \varepsilon$$

- $\mathbf{\Sigma} = \mathbf{\Gamma}\mathbf{\Omega}\mathbf{\Gamma}^\top + \mathbf{\Gamma}_0\mathbf{\Omega}_0\mathbf{\Gamma}_0^\top$ does not need anymore
- $\mathcal{S}(\mathbf{\Sigma}\mathbf{\Gamma}) \subseteq \mathcal{S}(\mathbf{\Gamma})$ iff $\mathcal{S}(\mathbf{\Sigma}_y\mathbf{\Gamma}) \subseteq \mathcal{S}(\mathbf{\Gamma})$
  $\Rightarrow E(\mathbf{Y}|\mathbf{X}) = E(\mathbf{P}_{\mathbf{\Gamma}(\mathbf{\Sigma}_y)}^\top\mathbf{Y}|\mathbf{X})$ if $\mathcal{S}(\mathbf{\Gamma})$ is invariant for $\mathbf{\Sigma}_y$

---

- $\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Sigma}}_{\text{res}} + \hat{\mathbf{\Sigma}}_{\text{res}}^{1/2}\hat{\mathbf{V}}\hat{\mathbf{K}}_d\hat{\mathbf{V}}^\top\hat{\mathbf{\Sigma}}_{\text{res}}^{1/2} = \hat{\mathbf{\Sigma}}_{\text{res}}^{1/2}(\mathbf{I}_r + \hat{\mathbf{V}}\hat{\mathbf{K}}_d\hat{\mathbf{V}}^\top)\hat{\mathbf{\Sigma}}_{\text{res}}^{1/2}$
- $\hat{\mathcal{S}}(\mathbf{\Gamma}) = \hat{\mathbf{\Sigma}}^{1/2}\mathcal{S}_d(\mathbf{B}) = \hat{\mathbf{\Sigma}}_{\text{res}}^{1/2}\mathcal{S}_d(\mathbf{B}_{\text{res}})$

  - $\mathbf{B}_d$: 1st $d$ eigenvectors of $\mathbf{B}$
  - $\mathbf{B} = \hat{\mathbf{\Sigma}}^{-1/2}\hat{\mathbf{\Sigma}}_{\text{fit}}\hat{\mathbf{\Sigma}}^{-1/2}$
  - $\hat{\mathbf{\Lambda}} = (\hat{\lambda}_1, ..., \hat{\lambda}_q)$: Eigenvalues of $\mathbf{B}_{\text{res}}$
  - $\hat{\mathbf{K}}_d = \operatorname{diag}(0, ..., 0, \hat{\lambda}_{d+1}, ..., \hat{\lambda}_q)$
  - $\mathcal{S}_d(\mathbf{B})$: Column subspace of $\mathbf{B}_d$
  - $\mathbf{B}_{\text{res}} = \hat{\mathbf{\Sigma}}_{\text{res}}^{-1/2}\hat{\mathbf{\Sigma}}_{\text{fit}}\hat{\mathbf{\Sigma}}_{\text{res}}^{-1/2}$
  - $\hat{\mathbf{V}} = (\hat{\gamma}_1, ..., \hat{\gamma}_q)$: Eigenvectors of $\mathbf{B}_{\text{res}}$

  $\Rightarrow$ Inverse of $\mathbf{\Sigma}_{\text{res}}$ and $\mathbb{F}^\top\mathbb{F}$ (from $\mathbf{\Sigma}_{\text{fit}}$) is needed

# Implementation along with Sample Size

*p*: **dimension of predictor**, *r*: **dimension of response**, *n*: **sample size**

- **YC-method** (Non-parametric method)
  - Requires inverse matrices of $\hat{\boldsymbol{\Sigma}}_x$ and $\hat{\boldsymbol{\Sigma}}_y$
    
    $\Rightarrow p < n, r < n$
- **PRR** (Principal Response Reduction)
  - Does not require inverse matrix
    
    $\Rightarrow$ No constraints for *p*, *r*, and *n*
- **PFRR** (Principal Fitted Response Reduction)
  - Requires inverse matrix of $\mathbb{F}^{\mathsf{T}}\mathbb{F}$
    
    $\Rightarrow p < n$
- **UPFRR** (Unstructured Principal Fitted Response Reduction)
  - Requires inverse matrices of $\mathbb{F}^{\mathsf{T}}\mathbb{F}$ and $\hat{\boldsymbol{\Sigma}}_{\text{res}}$
    
    $\Rightarrow p < n, r < n$

# Implementation along with Sample Size

**Table 1. Feasible response dimension reduction methods** according to $p$, $r$ and $n$

|        | Relation between *p, r, n* | Feasible methods |
|--------|:--------------------------:|:----------------:|
| **case 1** | $\max(p, r) < n$ | YC-method, PRR, PFRR, UPFRR |
| **case 2** | $p \geq n$, but $r < n$ | PRR |
| **case 3** | $p < n$, but $r \geq n$ | PRR, PFRR |
| **case 4** | $\min(p, r) \geq n$ | PRR |

- **case 2**: $r < n \leq p$
  - Reducing $p$ with **multivariate seeded** dimension reduction (Yoo and Im, 2014)
    $\Rightarrow$ Enable YC-method, PFRR, UPFRR
- **case 3**: $p < n \leq r$
  - Reducing $r$ through **PRR** or **PFRR**
    $\Rightarrow$ Enable YC-method, UPFRR
- **case 4**: $n \leq \min(p, r)$
  1. Reducing $r$ through **PRR**
  2. Reducing $p$ with **multivariate seeded** dimension reduction
     $\Rightarrow$ Enable YC-method, PFRR, UPFRR

# Outline

## Real data analysis I (case 2: $p \geq n$, $r < n$)

**Near-infrared (NIR) Spectroscopy of Biscuit Doughs data**

- `cookie` dataset from the **seedCCA** package (Kim et al., 2021)
    - **X** ($p = 256$): **wavelengths** spanning from 1380 to 2400nm at 4nm intervals
    - **Y** ($r = 4$): **fat** (*Y*1), **sucrose** (*Y*2), **dry flour** (*Y*3), **water** (*Y*4)
    - $n_{\text{train}} = 39$, $n_{\text{test}} = 31$ ($n \leq p$) $\Rightarrow$ Require reducing *p*

  $\Rightarrow$ **X**$_{\text{train}}$: 39 $\times$ 256, **X**$_{\text{test}}$: 31 $\times$ 256, **Y**$_{\text{train}}$: 39 $\times$ 4, **Y**$_{\text{test}}$: 31 $\times$ 4

- Based on the **multivariate seeded** dimension reduction,
  **X** ($p = 256$) $\rightarrow$ **Rx** ($p' = 4$)
    - **Rx**$_{\text{train}}$: 39 $\times$ 4, **Rx**$_{\text{test}}$: 31 $\times$ 4 ($p' < n$)
  $\Rightarrow$ **UPFRR**, **PFRR**, and **PRR** are applied

**Table 2. UPFRR and PFRR Results for `cookie` data**

| Comparison | Stat | df | p-value | Comparison | Stat | df | p-value |
|---|---|---|---|---|---|---|---|
| **0D vs $\geq$ 1D** | 340.9 | 32 | <**0.0001** | **0D vs $\geq$ 1D** | 340.9 | 32 | <**0.0001** |
| **1D vs $\geq$ 2D** | 159.4 | 21 | <**0.0001** | **1D vs $\geq$ 2D** | 196.1 | 24 | <**0.0001** |
| **2D vs $\geq$ 3D** | 27.39 | 12 | **0.0068** | **2D vs $\geq$ 3D** | 63.80 | 16 | <**0.0001** |
| 3D vs $\geq$ 4D | 9.508 | 5 | 0.0904 | 3D vs $\geq$ 4D | 11.38 | 8 | 0.1809 |

(a) **UPFRR** Test Summary                    (b) **PFRR** Test Summary

- For **(Rx, Y)**, both **UPFRR** and **PFRR** suggest reducing **Y** to 3 dimensions
  (5% significance level)

**Table 3. Cumulative Variance** of 1-4 Dimensions with Each RDR Method

| Method | 1D | 2D | 3D | 4D |
|---|---|---|---|---|
| **UPFRR** | 0.1136 | 0.3470 | 0.6676 | 1.0000 |
| **PFRR** | 0.1915 | 0.4281 | 0.7059 | 1.0000 |
| **PRR** | **0.7085** | **0.9902** | 1.0000 | 1.0000 |

- **PRR** shows that 1st 2 components account for 99.02% of variance
  $\Rightarrow$ **UPFRR** and **PFRR** may overestimate the number of dimensions

# Real data analysis I (case 2: $p \geq n$, $r < n$)

## Trace correlation coefficients (Hooper, 1959)

Given two matrices $A_1$ and $A_2$, $q$ is the number of principal components to consider:

$$\bar{r} = \sqrt{\frac{\sum_{i=1}^{q} \lambda_i}{q}}$$

where $\lambda_i$ are the eigenvalues of the matrix $P = A_1^{\mathsf{T}} A_2 A_2^{\mathsf{T}} A_1$.

**Table 4.** Trace Correlation Coefficients for **UPFRR**, **PFRR**, and **PRR**

| Comparison | 1st 2 components | 1st 3 components |
|:---:|:---:|:---:|
| **UPFRR vs PFRR** | 0.9247 | 1.0000 |
| **UPFRR vs PRR** | 0.9218 | 1.0000 |
| **PFRR vs PRR** | 1.0000 | 1.0000 |

- Similarity in the 1st 2 components suggests 2-dimensional **Ry** is sufficient
  $\Rightarrow$ Apply **PFRR** to reduce **Y** ($r = 4$) $\rightarrow$ **Ry** ($r' = 2$)

# Real data analysis I (case 2: $p \geq n$, $r < n$)

- In **Ry** ($r' = 2$),
  - **1st component**: linear combination of sucrose (*Y2*), dry flour (*Y3*), water (*Y4*)
  - **2nd component**: fat (*Y1*)



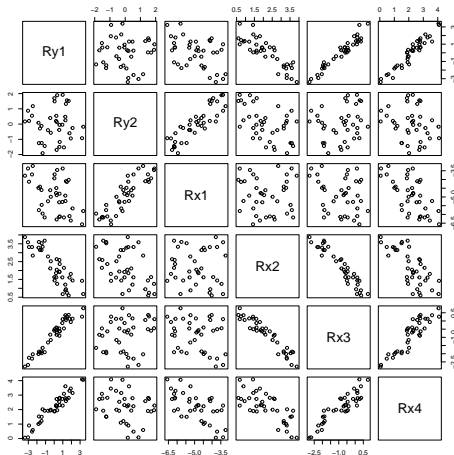**Figure 1.** Scatter plot matrix of **(Ry, Rx)** for `cookie` data

# Real data analysis I (case 2: $p \geq n$, $r < n$)

- **(Ry, Rx)**, **(Y, Rx)**: Analyze using **multivariate linear regression**
- **(Ry, X)**, **(Y, X)**: Analyze using **partial least squares regression**
  - **mvr** function from **pls** package (Liland et al., 2021)

- **Mean Squared Prediction Error** (MSPE)

$$\text{MSPE} = \frac{1}{n \times r} \sum_{i=1}^{n} \sum_{j=1}^{r} (y_{ij} - \hat{y}_{ij})^2$$

**Table 5. MSPE** Results for **cookie** data

| Model | *Y1* | *Y2* | *Y3* | *Y4* |
|---------|--------|--------|--------|--------|
| **(Ry, Rx)** | 0.0901 | 0.0462 | 0.0452 | 0.0471 |
| **(Y, Rx)** | 0.0937 | 0.0463 | 0.0506 | 0.0470 |
| **(Ry, X)** | 0.0397 | 0.0418 | 0.0462 | 0.0400 |
| **(Y, X)** | 0.0396 | 0.0418 | 0.0495 | 0.0392 |

- Comparing **(Ry, Rx)** vs **(Y, Rx)** and **(Ry, X)** vs **(Y, X)**
  $\Rightarrow$ No substantial difference in prediction performance

**Nutrimouse data**

- **nutrimouse** dataset from the **seedCCA** package (Kim et al., 2021)
    - $X_1$: **genotypes** - wild type, PPARalpha
    - $X_2$: **diets** - COC(coconut oil), FISH(fish oil), LIN(linseed oil),
              REF(corn and colza oil), SUN(sunflower oil)
    - **Y** ($r = 141$): 120 genes and 21 lipids
    - $n = 40$ ($n < r$) ⇒ Require reducing $r$
    - Predictors are **categorical**, **PRR** is the preferable choice

**Table 6. Cumulative Variance** of 1-6 Dimensions with **PRR** Method for **nutrimouse** data
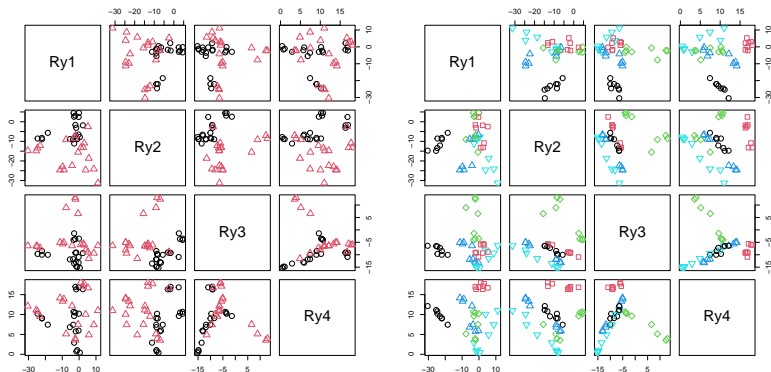
| Dimension | 1D | 2D | 3D | 4D | 5D | 6D |
|-----------|--------|--------|--------|--------|--------|--------|
| Variance | 0.4038 | 0.6949 | 0.8623 | 0.9554 | 0.9726 | 0.9835 |

- Based on **PRR**, 1st 4 components capture over 95% of total variance
  **Y** ($r = 141$) → **Ry** ($r' = 4$)

**Figure 2.** Scatter plot matrices of the reduced responses marked by **genotype** and **diet**
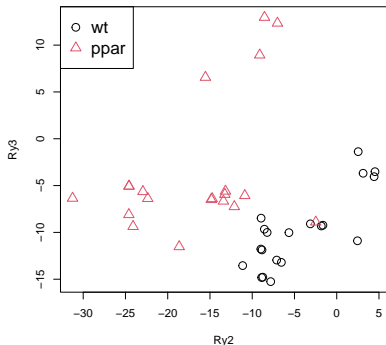


(a) Scatter plot matrix marked by **genotype**   (b) Scatter plot matrix marked by **diet**

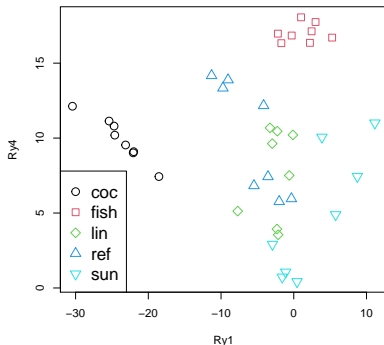- **Ry2** & **Ry3** distinctly separate **genotypes**, **Ry1** & **Ry4** distinctly separate **diets**

**Figure 3.** Scatter plots of **Ry2** and **Ry3** marked by genotype and of **Ry1** and **Ry4** marked by diet



(a) Plot of **Ry2** and **Ry3** (**genotype**)

(b) Plot of **Ry1** and **Ry4** (**diet**)

- **Ry2** & **Ry3** associated with **genotype**, **Ry1** & **Ry4** associated with **diet**
  ⇒ possible existence of interaction between **genotype** and **diet**

**Table 7. MANOVA** Results for `nutrimouse` data

| Effect | df | Pillai's trace | Approx. F | Pr(>F) |
|--------|-----|----------------|-----------|--------|
| **Genotype** | 1 | 0.9070 | 65.825 | **1.574e-13 \*\*\*** |
| **Diet** | 4 | 3.0805 | 25.125 | < **2.2e-16 \*\*\*** |
| **Genotype:Diet** | 4 | 2.0993 | 8.284 | **4.339e-13 \*\*\*** |
| **Residuals** | 30 | - | - | - |

- **MANOVA** with **Ry** shows that the effects of genotype, diet and their interaction are significant at 5%

# Real data analysis II (case 3: $p < n$, $r \geq n$)

- Split the data according to **genotype** (wild type, PPARalpha)
- Compare dietary **factor-level means** using `test` function in **lsmeans** package (Lenth, 2016)

**Table 8.** Pairwise Comparison Results for `nutrimouse` data

| Contrast | Estimate | df | t-ratio | p-value | Contrast | Estimate | df | t-ratio | p-value |
|----------|----------|-----|---------|---------|----------|----------|-----|---------|---------|
| coc-fish | -8.935 | 15 | -23.833 | <.0001 | coc-fish | -9.817 | 15 | -15.126 | <.0001 |
| coc-lin | -9.714 | 15 | -25.912 | <.0001 | coc-lin | -9.370 | 15 | -14.437 | <.0001 |
| coc-ref | -3.401 | 15 | -9.072 | <.0001 | coc-ref | -2.794 | 15 | -4.305 | 0.0008 |
| coc-sun | -1.627 | 15 | -4.341 | 0.0006 | coc-sun | -4.435 | 15 | -6.833 | <.0001 |
| fish-lin | -0.779 | 15 | -2.079 | 0.0552 | fish-lin | 0.447 | 15 | 0.689 | 0.5013 |
| fish-ref | 5.534 | 15 | 14.762 | <.0001 | fish-ref | 7.023 | 15 | 10.821 | <.0001 |
| fish-sun | 7.308 | 15 | 19.492 | <.0001 | fish-sun | 5.382 | 15 | 8.293 | <.0001 |
| lin-ref | 6.313 | 15 | 16.841 | <.0001 | lin-ref | 6.576 | 15 | 10.132 | <.0001 |
| lin-sun | 8.087 | 15 | 21.571 | <.0001 | lin-sun | 4.935 | 15 | <.0001 | <.0001 |
| ref-sun | 1.774 | 15 | 4.731 | 0.0003 | ref-sun | -1.641 | 15 | 0.0258 | 0.0258 |

(a) Pairwise Comparison in **Wild Type**     (b) Pairwise Comparison in **PPARalpha**

- In **Wild Type**, all diets are different from each other
- In **PPARalpha**, **FISH** and **LIN** are not distinguishable

# Summary and Limitations

- **Summary**
    - Demonstrate the practical application of **Response Dimension Reduction** across different relationships between *p*, *r*, *n*
    - Apply to `cookie` data (**case 2**: $n \leq p$), `nutrimouse` data (**case 3**: $n \leq r$)
      - Achieve simplified information while maintaining predictive performance

- **Limitations**
    - Beyond **case 1**, it is hard to capture **nonlinear relationship**
    - Absence of real data applications for **case 4**: $n \leq \min(p, r)$
    - Variable performance of RDR methods across different datasets

Thank You.

Hooper, J. W. (1959). Simultaneous equations and canonical correlation theory. *Econometrica: Journal of the Econometric Society*, 27(2):245–256.

Kim, B., Im, Y., and Yoo, J. K. (2021). SEEDCCA: An integrated R-package for canonical correlation analysis and partial least squares. *The R Journal*, 13(1):7–20.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, 69(1):1–33.

Liland, K. H., Mevik, B.-H., Wehrens, R., and Hiemstr, P. (2021). PLS: Partial least squares and principal component regression. R package version 2.8.0.

Yoo, J. K. (2018). Response dimension reduction: model-based approach. *Statistics*, 52(2):409–425.

Yoo, J. K. (2019). Unstructured principal fitted response reduction in multivariate regression. *Journal of the Korean Statistical Society*, 48(4):561–567.

Yoo, J. K. and Cook, R. D. (2008). Response dimension reduction for the conditional mean in multivariate regression. *Computational Statistics and Data Analysis*, 53(2):334–343.

Yoo, J. K. and Im, Y. (2014). Multivariate seeded dimension reduction. *Journal of the Korean Statistical Society*, 43(4):559–566.