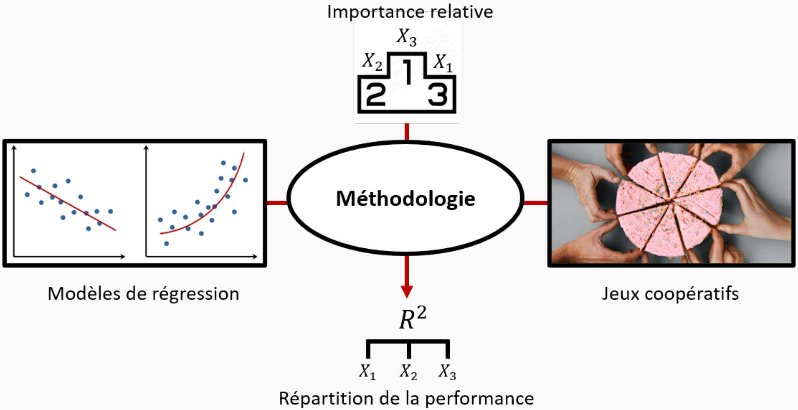
Quantification de l'importance relative et jeux coopératifs

EDF R&D - *Département PRISME*
SINCLAIR - *AI Lab.*

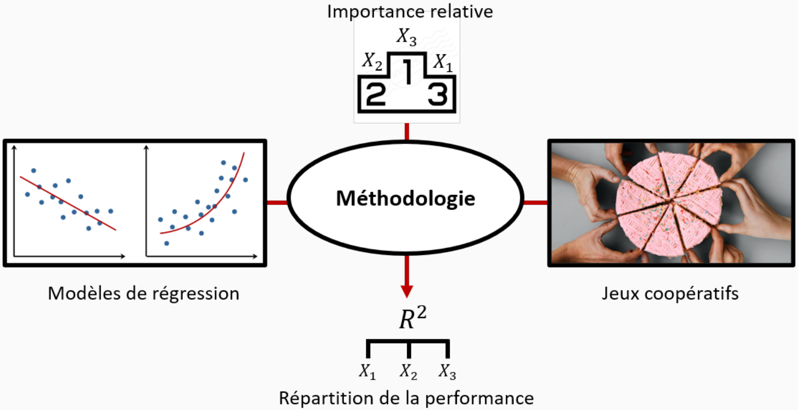
Séminaire PSPP
10 Février 2021

Marouane Il Idrissi

Introduction



Introduction



Comment construire des mesures d'importance relative tirées de la théorie des jeux coopératifs ?

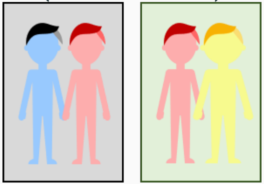
Sommaire

Sommaire

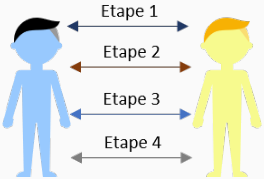
Jeux coopératifs

Deux domaines d'étude de la théorie des jeux :

Jeux coopératifs
(combinatoires)



Jeux non-coopératifs
(procéduraux)



Modélisation de l'issue d'un jeu lorsque les joueurs s'organisent en coalitions différentes.

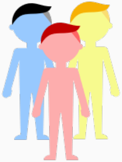
Modélisation des décisions offertes aux joueurs de manière procédurale.

Les jeux coopératifs ne traitent pas exclusivement de coopération (et inversement).

Jeux coopératifs

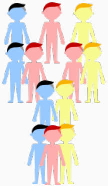
Joueurs

N



Coalitions

$\mathcal{P}_n$



Fonction de coût

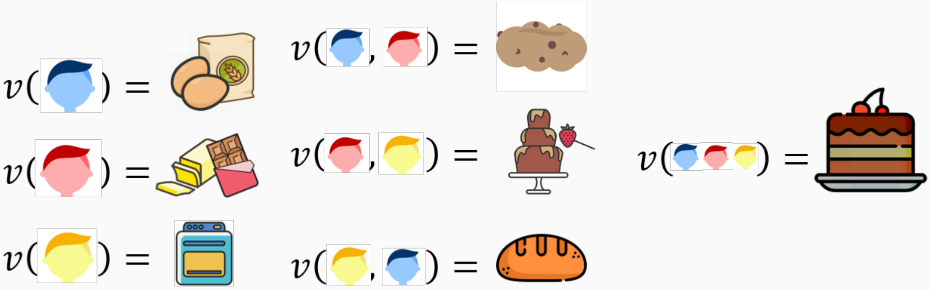
v

Quantifie la valeur produite par une coalition

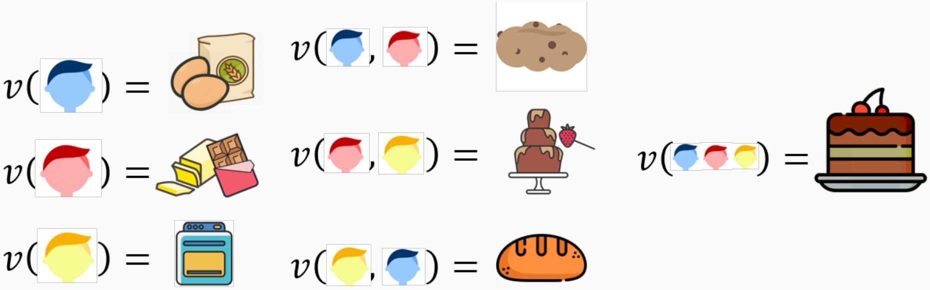
Soit $N = \{1, \ldots, n\}$ un ensemble fini dénombrable de joueurs, et $\mathcal{P}_n$ l'ensemble des sous-ensembles de N (coalitions). Soit v une fonction de coût de $\mathcal{P}_n$ dans $\mathbb{R}$.

Le jeu coopératif composé des joueurs N, et de la fonction de coût v est formellement défini par le tuple (N, v).

Allocations



Allocations



Comment allouer une part de la valeur totale produite à chacun des joueurs ?

Allocations

Une allocation relative à un jeu (N, v) est une collection $\phi = (\phi_1, \dots, \phi_n) \in \mathbb{R}^n$.

On peut imaginer un **processus de négociation** entre les joueurs, pour déterminer une allocation.

Les **contributions marginales** permettent de quantifier le **pouvoir de négociation** d'un joueur ou d'une coalition. Soit $S \in \mathcal{P}_n$, la contribution marginale de S est notée :

$$w(S) = v(N) - v(N \setminus S). \tag{1}$$

Allocations

Une allocation relative à un jeu (N, v) est une collection $\phi = (\phi_1, \ldots, \phi_n) \in \mathbb{R}^n$.

On peut imaginer un **processus de négociation** entre les joueurs, pour déterminer une allocation.

Les **contributions marginales** permettent de quantifier le **pouvoir de négociation** d'un joueur ou d'une coalition. Soit $S \in \mathcal{P}_n$, la contribution marginale de S est notée :

$$w(S) = v(N) - v(N \setminus S). \tag{1}$$

Exemples de propriétés désirées :

- **Efficacité :** $\sum_{i=1}^{n} \phi_i = v(N)$.
- **Individuellement rationnelle :** $\phi_i \geq v(\{i\}), \forall i \in N$.
- **Symétrie :** Si $\forall S \in \mathcal{P}_n$, $v(S \cup \{i\}) = v(S \cup \{j\})$, $i \neq j$ alors $\phi_i = \phi_j$.

Modèles à ordres aléatoires

Une manière de construire une allocation est de considérer les **modèles à ordres aléatoires** (Feldman 2007).

Idée :

- Considérer que les joueurs **sont mis en relation dans un certain ordre** (une permutation de N), considéré aléatoire.
- Chaque joueur reçoit **sa contribution marginale au groupe de joueurs le précédant**, pondérée par la **"vraisemblance" de cet ordre**.

Soit $\mathcal{R}(N)$ l'ensemble des $n!$ ordres/permutations de N. Par exemple, si $N = \{1, 2, 3\}$, alors :

$$\mathcal{R}(N) = \Big\{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\Big\}. \qquad (2)$$

Modèles à ordres aléatoires

Soit $r = (r_1, \dots, r_n) \in \mathcal{R}(N)$ une permutation de N.
L'ensemble $S_k^r = \{r_1, \dots, r_k\}$ contient les k premiers éléments de l'ordre r.
La **contribution marginale individuelle positionnelle** du $i^{\text{ème}}$ joueur de r est notée :

$$\begin{aligned} M_i(r) &= w(S_i^r) - w(S_{i-1}^r) && (3) \\ &= v(N \setminus S_{i-1}^r) - v(N \setminus S_i^r) && (4) \end{aligned}$$

Et peut être interprétée comme la **contribution marginale de** r_i **à** $v(N \setminus \{r_1, \dots, r_{i-1}\})$.

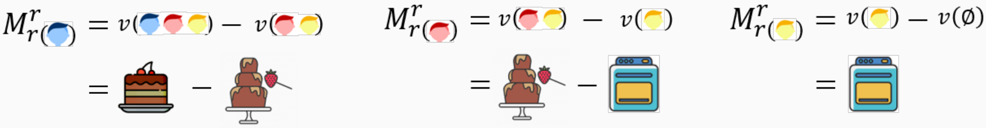
Notons :

- $r(i)$ la position du joueur $i \in N$ dans la permutation r, de sorte que $r_{r(i)} = i$.
- $p(r)$ une fonction de masse définie sur $\mathcal{R}(N)$.

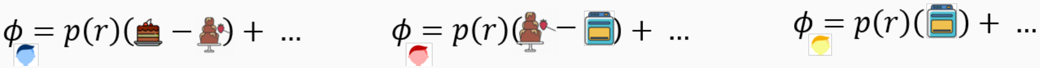
Modèles à ordres aléatoires

Ordre aléatoire

Contributions marginales individuelles positionnelles



Allocation par modèle à ordre aléatoire



Allocation par modèle à ordre aléatoire

Alors, pour $i = 1, \ldots, n$:

$$\begin{aligned}
\phi_i &= \mathbb{E}_p\big[M_i(r)\big] && (5)\\
&= \sum_{r \in \mathcal{R}(N)} p(r) M_{r(i)}(r) && (6)\\
&= \sum_{r \in \mathcal{R}(N)} p(r)\Big(w(S^r_{r(i)}) - w(S^r_{r(i)-1})\Big) && (7)\\
&= \sum_{r \in \mathcal{R}(N)} p(r)\Big(v(N \setminus S^r_{r(i)-1}) - v(N \setminus S^r_{r(i)})\Big) && (8)
\end{aligned}$$

est l'allocation par modèle à ordre aléatoire de (N, v).

Propriété (Efficacité (Feldman 2005))

Pour toute fonction de masse p définie sur $\mathcal{R}(N)$, l'allocation ϕ par modèle à ordre aléatoire d'un jeu coopératif (N, v) est efficace, i.e., :

$$\sum_{i=1}^{n} \phi_i = v(N).$$

Jeux coopératifs statistiques

Soit $X = (X_1, \ldots, X_n)$ les covariables d'un modèle paramétrique Θ "emboîtable", tel que :

$$Y = \Theta(X, \beta) \tag{9}$$

avec $\beta \in \mathbb{R}^n$.

Pour $S \in \mathcal{P}_n$, on note $\Theta(X_S, \beta_S)$ le sous-modèle emboîté de $\Theta(X, \beta)$. Par exemple :

$$Y = \beta_1 X_1 + \beta_2 X_2 + \epsilon \tag{10}$$

est emboîté dans le **modèle total** :

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \tag{11}$$

Jeux coopératifs statistiques

Soit $\mu_\Theta(S) : \mathcal{P}_n \to \mathbb{R}^+$ une **mesure de performance** de $\Theta(X_S, \beta_S)$, pour tout $S \in \mathcal{P}_n$.

Elle est supposée **faiblement monotone** : un sous-modèle ne peut pas être plus performant qu'un modèle plus grand.

Pour un modèle linéaire, μ_Θ peut être le R^2 (coefficient de détermination), la **variance expliquée**, ou la **(log-)vraisemblance (négative)** du modèle Θ.

Jeux coopératifs statistiques

Joueurs	**Coalitions**	**Fonction de coût**
N	$\mathcal{P}_n$	$\mu_\Theta(S)$
$X_1, \dots, X_n$	$\{X_1\}, \{X_2\}, \dots,$ $\{X_1, X_2\}, \{X_1, X_3\}, \dots,$ $\{X_1, \dots, X_n\}$	*Performance du sous-modèle* $\Theta(X_S, \beta_S)$

Un jeu coopératif statistique est le jeu coopératif défini par (N, μ_Θ).

Propriété (Efficacité, positivité (Feldman 2005))

Pour tout p, l'allocation ϕ par modèle à ordre aléatoire d'un jeu coopératif statistique (N, μ_Θ) est efficace *et* positive, *i.e., :*

$$\sum_{i=1}^{n} \phi_i = \mu_\Theta(N), \quad \phi_i \geq 0, \forall i \in N.$$

ϕ est donc une **décomposition** de $\mu_\Theta(N)$ selon p.

Jeux coopératifs statistiques

Joueurs N	Coalitions $\mathcal{P}_n$	Fonction de coût $\mu_\Theta(S)$
$X_1, \dots, X_n$	$\{X_1\}, \{X_2\}, \dots,$ $\{X_1, X_2\}, \{X_1, X_3\}, \dots,$ $\{X_1, \dots, X_n\}$	*Performance du sous-modèle* $\Theta(X_S, \beta_S)$

Un jeu coopératif statistique est le jeu coopératif défini par (N, μ_Θ).

Propriété (Efficacité, positivité (Feldman 2005))

Pour tout p, l'allocation ϕ par modèle à ordre aléatoire d'un jeu coopératif statistique (N, μ_Θ) est efficace *et* positive, *i.e., :*

$$\sum_{i=1}^{n} \phi_i = \mu_\Theta(N), \quad \phi_i \geq 0, \forall i \in N.$$

ϕ est donc une **décomposition** de $\mu_\Theta(N)$ selon p.

Comment définir p ?

Modélisation de l'importance relative

L'**importance** relative peut être considérée comme une **relation d'ordre** (complète et transitive) entre les covariables (Feldman 2005) dans un modèle.

Soit $\prec$ une relation d'ordre binaire définie sur N **décrivant l'importance relative** de covariables pour un modèle $\Theta(X, \beta)$, avec mesure de performance μ_Θ :

- $i \prec j$ signifie que X_j est *plus important* que X_i.
- $i \sim j$ signifie que X_i et X_j sont d'importance équivalente.

Définition (Représentation d'un ordre binaire)

Soit $\prec$ une relation d'ordre binaire définie sur un ensemble N. Une fonction $\phi : N \to \mathbb{R}$ est une représentation (utilité) de $\prec$ si, pour tout $i, j \in N$:

$$i \preceq j \iff \phi_i \leq \phi_j.$$

Une **mesure d'importance relative** est donc **une représentation** de $\prec$.

Modélisation de l'importance relative

L'existence de représentations de $\prec$ est garantie (Kreps 2013).

Idée : faire en sorte que p **mesure la "vraisemblance" qu'un ordre** r **respecte** $\prec$.

i.e., si $r_1 \preceq r_2 \preceq \cdots \preceq r_n$, alors $p(r)$ doit être grand.

Permet de produire des **allocations candidates à représenter** $\prec$.

Modélisation de l'importance relative

L'existence de représentations de $\prec$ est garantie (Kreps 2013).

Idée : faire en sorte que p **mesure la "vraisemblance" qu'un ordre** r **respecte** $\prec$.

i.e., si $r_1 \preceq r_2 \preceq \cdots \preceq r_n$, alors $p(r)$ doit être grand.

Permet de produire des **allocations candidates à représenter** $\prec$.

L'importance relative $\prec$ est **inconnue**, **inobservée**, et définie de manière très générale.

Admissibilité d'une mesure d'importance relative

Comment définir une **mesure d'importance relative admissible** ?

Quatre critères d'admissibilité (Cox 1985 ; Johnson et Lebreton 2004 ; Feldman 2005 ; Grömping 2007) :

- **Positivité :** $\forall i \in N, \phi(i) \geq 0$.
- **Exclusion :** Si, pour $\Theta(X, \beta)$, $\beta_i = 0$, alors $\phi_i = 0$.
- **Inclusion :** Si, pour $\Theta(X, \beta)$, $\beta_i \neq 0$, alors $\phi_i > 0$.
- **Contribution totale :** $\sum_{i=1}^{n} \phi_i = \mu_\Theta(N)$.

Méthodologie de construction de mesures de l'importance relative

Pour résumer :

1. On prend un **jeu coopératif statistique** (N, μ_Θ).
2. On définit une **allocation par modèle aléatoire** ϕ sur (N, μ_Θ).
3. On choisit p afin que ϕ soit **candidat à représenter** l'importance relative $\prec$ sur le modèle Θ.
4. On étudie les **propriétés d'admissibilité** de l'allocation ϕ.

Sommaire

Valeur de Shapley

Concept introduit par Shapley (1951) dans le contexte des jeux coopératifs.

L'**unique** allocation qui satisfait, pour un jeu coopératif (N, v) :

1. **Efficacité :** $\sum_{i=1}^{n} \mathrm{Shap}_i = v(N)$.
2. **Symétrie :** si $\forall S \in \mathcal{P}_n$, $v(S \cup \{i\}) = v(S \cup \{j\})$, $i \neq j$ alors $\phi_i = \phi_j$.
3. **Joueur nul :** si $v(S \cup \{i\}) = 0, \forall S \in \mathcal{P}_n$, alors $\mathrm{Shap}_i = 0$.
4. **Additivité :** soient deux jeux (N, v) et (N, v') ayant pour allocation ϕ et ϕ', alors le jeu $(N, v + v')$ a pour allocation $\phi + \phi'$.

Valeur de Shapley

Dans le contexte des **modèles à ordres aléatoires** pour un jeu coopératif statistique (N, μ_Θ), pour tout $r \in \mathcal{R}(N)$:

$$p(r) = \frac{1}{n!}$$

Ce qui donne, $i = 1, \ldots, n$:

$$\mathrm{Shap}_i = \frac{1}{n!} \sum_{r \in \mathcal{R}(N)} M_{r(i)}(r)$$

- **Les valeurs de Shapley de** (N, μ_Θ) **ne respectent pas le critère d'exclusion** (Feldman 2005).
- Peuvent être assimilé à un choix de **maximum d'entropie** (Soofi, Retzer et Yasai-Ardekani 2000) ou un **a priori uniforme** sur $\mathcal{R}(N)$.

Indices d'importance LMG

Dans le cadre d'un modèle Θ de **régression linéaire**, avec $\mu_\Theta = R^2$, i.e., :

$$R^2(S) = 1 - \frac{\mathbb{E}\left[(Y - \Theta(X_S, \beta_S))^2\right]}{\mathbb{V}(Y)} \tag{12}$$

L'allocation par **valeur de Shapley** du jeux coopératif statistique (N, R^2) sont les indices LMG d'après Lindeman, Merenda et Gold (1980).

Dans le cadre de **modèles plus généraux**, ils sont équivalent à la définition alternative des **effets de Shapley** (Owen 2014) introduite par Song, Nelson et Staum (2016), et étudiés extensivement par Iooss et Prieur (2019), avec $\mu_\Theta = E_S$:

$$E_S = \mathbb{E}\left[\mathbb{V}(Y \mid X_{N \setminus S})\right] \tag{13}$$

Décomposition marginale proportionnelle (PMD)

Inspirée de l'allocation par **valeur proportionnelle** des jeux coopératifs introduite par Ortmann (2000).

Déclinée aux **jeux coopératifs statistiques** par Feldman (2005).

μ_Θ peut donner **une idée** sur l'ordonnancement des covariables selon $\prec$. Si une permutation $r \in \mathcal{R}(N)$ satisfait :

$$M_1(r) < M_2(r) < \cdots < M_n(r)$$

alors, il y a de *fortes chances* que r respecte l'ordre d'importance relative (i.e., $r_1 \preceq r_2 \preceq \cdots \preceq r_n$).

Idée : tirer profit de l'information qu'offre μ_Θ sur $\prec$ dans la définition de p.

Décomposition marginale proportionnelle (PMD)

Posons :

$$p(r) = \frac{L(r)}{\sum_{m \in \mathcal{R}(N)} L(m)} = P(N) * L(r).$$

Notons $S \in r$ pour tout ensemble $\{r_1, \ldots, r_i\}$ tel quel $S \in \left\{S_i^r\right\}_{i=1}^n$.
Par exemple, pour $r = (2, 3, 1)$, on a que :

$$S \in r \implies S \in \left\{ \begin{matrix} \{2\}, \\ \{2,3\}, \\ \{2,3,1\} \end{matrix} \right\}. \tag{14}$$

Décomposition marginale proportionnelle (PMD)

Trois axiomes définissent l'allocation (PMD) (Feldman 2005) :

1. **Anonymité :** soient r et r^* deux permutations dans $\mathcal{R}(N)$. Si $w(S_i^r) = w(S_i^{r^*})$ pour tout $i \in N$ alors :
$$L(r) = L(r^*).$$
2. **Exclusion** (simplifié)[1] : si $w(\{i\}) = 0$ alors $\phi_i = 0$.
3. **Effets proportionnels égaux :** pour $r \in \mathcal{R}(N)$, et pour tout $S \in r$:
$$\left| \frac{\partial \ln L(r)}{\partial \ln w(S)} \right| = 1.$$

1. Une version limite plus technique de cet axiome est nécessaire.

Identification de la (PMD)

Selon Feldman (2005), l'**unique** fonction de masse respectant les trois axiomes est définie par :

$$L(r) = \left(\prod_{S \in r} w(S)\right)^{-1}$$

avec pour facteur de normalisation :

$$P(N) = \left(\sum_{r \in \mathcal{R}(N)} L(r)\right)^{-1} \tag{15}$$

$$= \left(\sum_{r \in \mathcal{R}(N)} \left(\prod_{S \in r} w(S)\right)^{-1}\right)^{-1}. \tag{16}$$

Allocation PMD

L'allocation ainsi obtenue est :

$$\phi_i(w) = \mathbb{E}_p\Big[M_i(r)\Big] \tag{17}$$

$$= P(N) \times \sum_{r \in \mathcal{R}(N)} L(r) M_{r(i)}(r) \tag{18}$$

L'allocation PMD du jeu coopératif statistique (N, μ_Θ) respecte **les quatre critères d'admissibilité** d'une mesure d'importance relative (Feldman 2005), tout en prenant en compte l'information contenue dans μ_Θ sur $\prec$.

Décomposition marginale proportionnelle de la variance (PMVD)

L'allocation PMD appliquée à un modèle Θ linéaire, avec mesure de performance $\mu_\Theta = R^2$ s'intitule **Proportional Marginal Variance Decomposition** (PMVD) (Feldman 2005).

Cette mesure d'importance relative a été extensivement étudiée dans Grömping (2007) et Grömping (2015).

La PMVD, dans le cadre d'un modèle linéaire, permet de **détecter des covariables non-influentes** en présence de **dépendance entre les entrées**.

Sommaire

Résultats analytiques

On considère le modèle :

$$Y = \beta_1 X_1 + \beta_2 X_2 \tag{19}$$

avec $\beta_1, \beta_2 \in \mathbb{R}$, et :

$$X \sim \mathcal{N}_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho \\ \sigma_1\sigma_2\rho & \sigma_2^2 \end{pmatrix} \right) \tag{20}$$

On prend R^2 comme mesure de performance. On a donc :

$$\mathbb{V}(Y) = \color{cyan}{\beta_1^2\sigma_1^2} + \color{red}{\beta_2^2\sigma_2^2} + \color{green}{2\beta_1\beta_2\sigma_1\sigma_2\rho} \tag{21}$$

Comme X est Gaussien, on a que :

$$\begin{aligned} R^2(\emptyset) &= 0 & R^2(\{1\}) &= (\beta_1^2\sigma_1^2 + 2\beta_1\beta_2\gamma + \beta_2^2\sigma_2^2\rho^2\ )/\mathbb{V}(Y) \\ R^2(\{2\}) &= (\beta_2^2\sigma_2^2 + 2\beta_1\beta_2\sigma_1\sigma_2\rho + \beta_1^2\sigma_1^2\rho^2\ )/\mathbb{V}(Y) & R^2(\{1,2\}) &= 1. \end{aligned} \tag{22}$$

Modèle à deux covariables

Dans le cas $n = 2$:

Valeurs de Shapley de (N, R^2) :

$$\mathrm{LMG}_1 = \frac{1}{\mathbb{V}(Y)}\left(\beta_1^2\sigma_1^2 + \beta_1\beta_2\sigma_1\sigma_2\rho + \frac{\rho^2}{2}(\beta_2^2\sigma_2^2 - \beta_1^2\sigma_1^2)\right)$$

$$\mathrm{LMG}_2 = \frac{1}{\mathbb{V}(Y)}\left(\beta_2^2\sigma_2^2 + \beta_1\beta_2\sigma_1\sigma_2\rho + \frac{\rho^2}{2}(\beta_1^2\sigma_1^2 - \beta_2^2\sigma_2^2)\right)$$

Valeurs proportionnelles de (N, R^2) :

$$\mathrm{PMVD}_1 = \frac{\beta_1^2\sigma_1^2}{\beta_1^2\sigma_1^2 + \beta_2^2\sigma_2^2}$$

$$\mathrm{PMVD}_2 = \frac{\beta_2^2\sigma_2^2}{\beta_1^2\sigma_1^2 + \beta_2^2\sigma_2^2}$$

L'allocation PMVD **ne dépend pas** de ρ (ne se généralise pas quand $n > 2$).

L'allocation LMG partage la part de variance due à la corrélation équitablement entre X_1 et X_2.

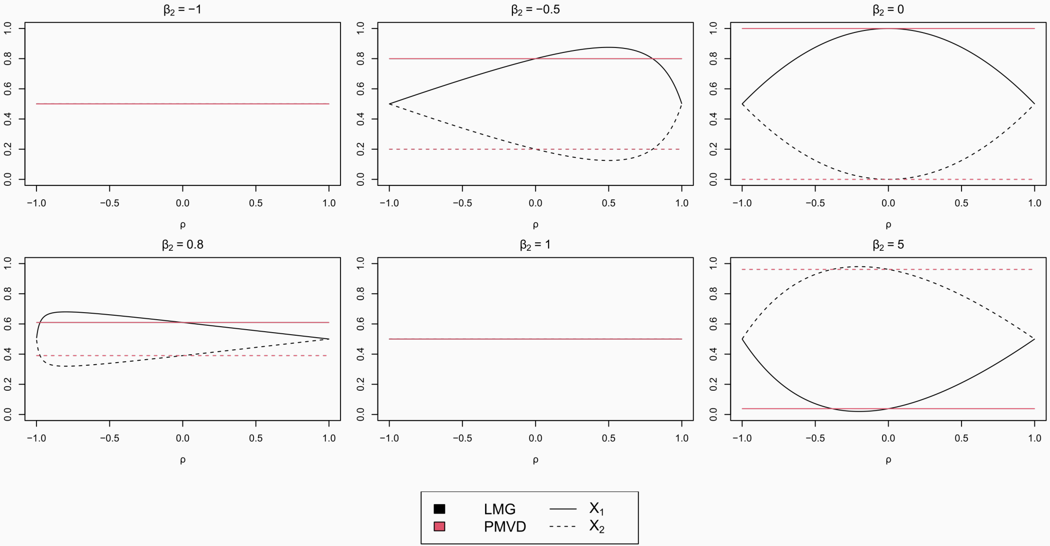


Figure 1: LMG et PMVD dans le cas bivarié pour $\beta_1 = 1$ et plusieurs valeurs de β_2, en fonction de ρ.

Modèle à trois covariables

On considère le modèle :

$$Y = X_1 + X_2 + \beta_3 X_3 \tag{23}$$

$\beta_3 \in \mathbb{R}$, et :

$$X \sim \mathcal{N}_3\left(\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1 & 0 & 0\\0 & 1 & \rho\\0 & \rho & 1\end{pmatrix}\right) \tag{24}$$

avec R^2 comme mesure de performance.

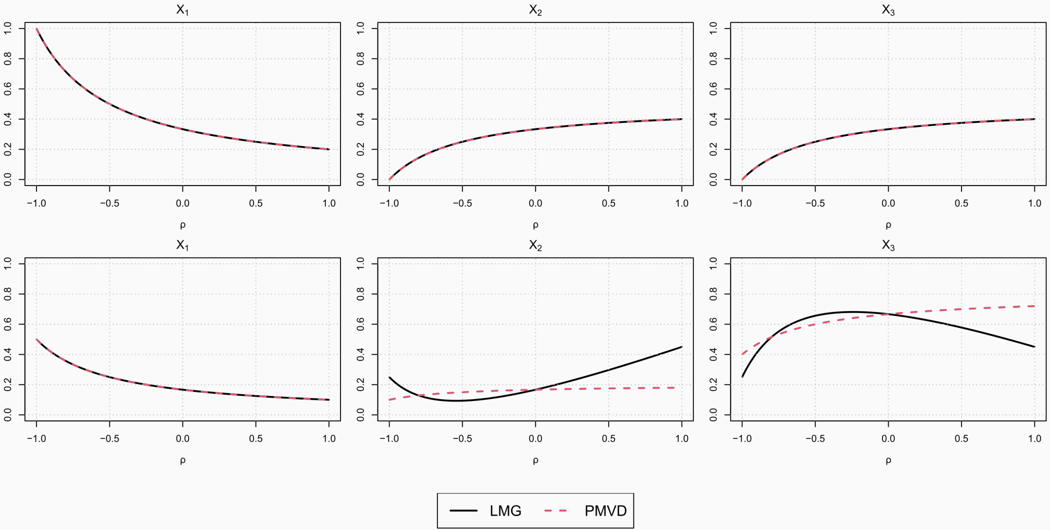


Figure 2: LMG et PMVD en fonction de ρ. Haut $\beta_3 = 1$, bas $\beta_3 = 2$.

Régression linéaire sur données "Airquality"

Données *Airquality* : 111 mesures complètes de 6 variables concernant la qualité de l'air à New York en 1973.

On chercher à prédire le niveau d'*Ozone* en fonction de 5 covariables :

- *Solar.R* : radiation solaire
- *Wind* : vitesse du vent moyenne
- *Temp* : température maximum
- *Month* : Mois de l'observation (entre Mai et Septembre)
- *Day* : Jour de l'observation

```r
library(sensitivity)

airq <- na.omit(airquality)
#Data types
airq$Month<-factor(airq$Month,
                   levels=5:9)
airq$Day<-factor(airq$Day,
                 levels=1:31)

#Relative importance measures
LMG.airq <- lmg(X=airq[,-1],
                y=airq$Ozone)
PMVD.airq <- emvd(X=airq[,-1],
                y=airq$Ozone)
```

Listing 1: R code Airquality.

Régression linéaire sur le dataset "Airquality"

Valeur de $R^2(N) = 0.816$.

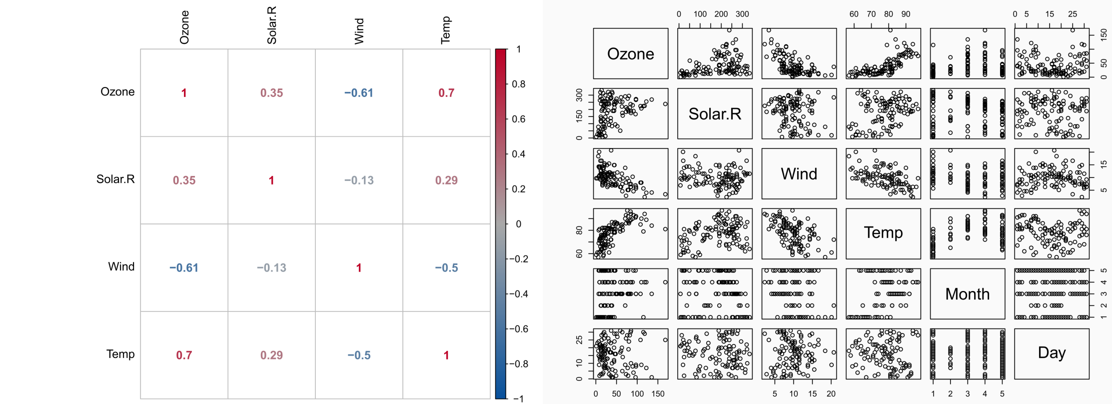


Figure 3: Corrélations (gauche) et nuages de points (droite) des données *Airquality*.

Régression linéaire sur données "Airquality"

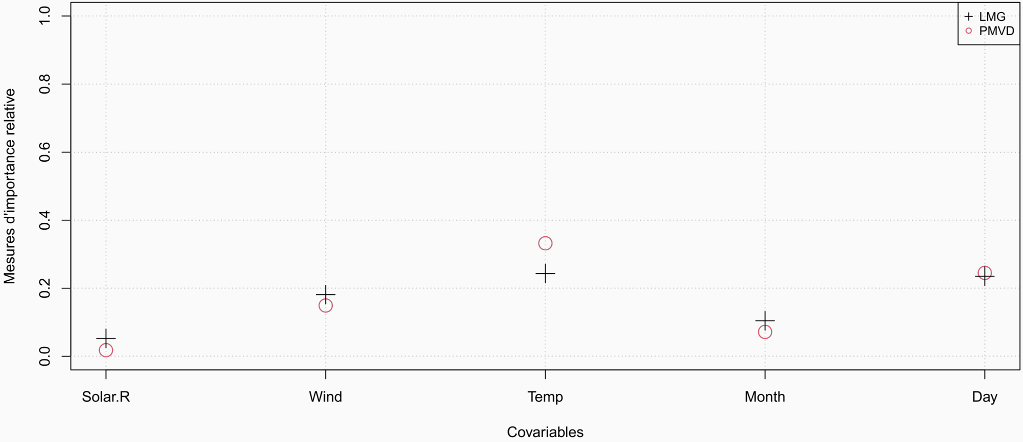


Figure 4: LMG et PMVD du modèle linéaire sur les données *Airquality*.

Conclusion récapitulative

- Formaliser le lien entre théorie des jeux coopératif et apprentissage statistique.
- Méthodologie de construction de mesures d'importance relative.
- L'allocation Shapley est un choix uniforme. La PMD est admissible.
- Résultats analytiques dans des cas gaussiens.
- Implémentation dans `sensitivity` et exemple d'utilisation.

Pistes d'ouverture

- Clarifier théoriquement le cas *régression logistique* :

$$R^2(S) = 1 - \frac{\text{Déviance de } \Theta(X_S, \beta_S)}{\text{Déviance modèle nul}}$$

- Étendre aux **modèles linéaires généralisés** et **modèles additifs généralisés**.
- Étendre à d'autres modèles "emboîtables" (e.g., arbres de décisions).
- PMD pour l'analyse de sensibilité.
- Autres définitions de p :
    - Explorer la décomposition d'autres mesures de performances μ_Θ.
    - p définie sur une autre quantité plus générale que μ_Θ (e.g., moment-independent indices).
    - Définitions de p différentes pour satisfaire d'autres jeux de critères

Cox, Louis A. 1985. "A New Measure of Attributable Risk for Public Health Applications" [en English]. *Management Science* 31, n$^{\circ}$ 7 (juillet) : 800-813. ISSN : 0025-1909, 1526-5501. https://doi.org/10.1287/mnsc.31.7.800. http://pubsonline.informs.org/doi/abs/10.1287/mnsc.31.7.800.

Feldman, B. E. 2007. "A Theory of Attribution" [en English]. *SSRN Electronic Journal*, ISSN : 1556-5068. https://doi.org/10.2139/ssrn.988860. http://www.ssrn.com/abstract=988860.

———. 2005. "Relative Importance and Value" [en English]. *SSRN Electronic Journal*, ISSN : 1556-5068. https://doi.org/10.2139/ssrn.2255827. http://www.ssrn.com/abstract=2255827.

Grömping, U. 2007. "Estimators of Relative Importance in Linear Regression Based on Variance Decomposition" [en English]. *The American Statistician* 61, n$^{\circ}$ 2 (mai) : 139-147. ISSN : 0003-1305, 1537-2731. https://doi.org/10.1198/000313007X188252. http://www.tandfonline.com/doi/abs/10.1198/000313007X188252.

———. 2015. "Variable importance in regression models" [en English]. *Wiley Interdisciplinary Reviews : Computational Statistics* 7, n$^{\circ}$ 2 (mars) : 137-152. ISSN : 19395108. https://doi.org/10.1002/wics.1346. http://doi.wiley.com/10.1002/wics.1346.

Iooss, B., et C. Prieur. 2019. "Shapley Effects For Sensitivity Analysis With Correlated Inputs : Comparisons With Sobol' Indices, Numerical Estimation And Applications" [en English]. *International Journal for Uncertainty Quantification* 9 (5) : 493-514. ISSN : 2152-5080. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2019028372. http://www.dl.begellhouse.com/journals/52034eb04b657aea,23ab8f375b210514,706e4963504bc249.html.

Références ii

Johnson, J. W., et J. M. Lebreton. 2004. "History and Use of Relative Importance Indices in Organizational Research" [en English]. *Organizational Research Methods* 7, n° 3 (juillet) : 238-257. ISSN : 1094-4281, 1552-7425. https://doi.org/10.1177/1094428104266510. http://journals.sagepub.com/doi/10.1177/1094428104266510.

Kreps, D. M. 2013. *Microeconomic foundations* [en English]. Princeton : Princeton University Press. ISBN : 978-0-691-15583-8.

Lindeman, R. H., P. F. Merenda et R. Z. Gold. 1980. *Introduction to Bivariate and Multivariate Analysis* [en English]. Scott, Foresman. ISBN : 978-0-673-15099-8. https://books.google.cz/books?id=-hfvAAAAMAAJ.

Ortmann, K. M. 2000. "The proportional value for positive cooperative games" [en English]. *Mathematical Methods of Operations Research* (*ZOR*) 51 (2) : 235-248. ISSN : 1432-2994, 1432-5217, visité le 4 décembre 2020. https://doi.org/10.1007/s001860050086. http://link.springer.com/10.1007/s001860050086.

Owen, A. B. 2014. "Sobol' Indices and Shapley Value" [en English]. *SIAM/ASA Journal on Uncertainty Quantification* 2, n° 1 (janvier) : 245-251. ISSN : 2166-2525, visité le 2 décembre 2020. https://doi.org/10.1137/130936233. http://epubs.siam.org/doi/10.1137/130936233.

Shapley, L. S. 1951. *Notes on the n-Person Game – II : The Value of an n-Person Game* [en English]. Research Memorandum ATI 210720. Santa Monica, California : RAND Corporation, août. https://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM670.pdf.

Song, E., B. L. Nelson et J. Staum. 2016. “Shapley Effects for Global Sensitivity Analysis : Theory and Computation” [en English]. *SIAM/ASA Journal on Uncertainty Quantification* 4, n° 1 (janvier) : 1060-1083. ISSN : 2166-2525. https://doi.org/10.1137/15M1048070. http://epubs.siam.org/doi/10.1137/15M1048070.

Soofi, E. S, J. J. Retzer et M. Yasai-Ardekani. 2000. “A Framework for Measuring the Importance of Variables with Applications to Management Research and Decision Models*” [en English]. *Decision Sciences* 31 (3) : 595-625. https://doi.org/https://doi.org/10.1111/j.1540-5915.2000.tb00936.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-5915.2000.tb00936.x.

Merci pour votre attention !

Des questions ?

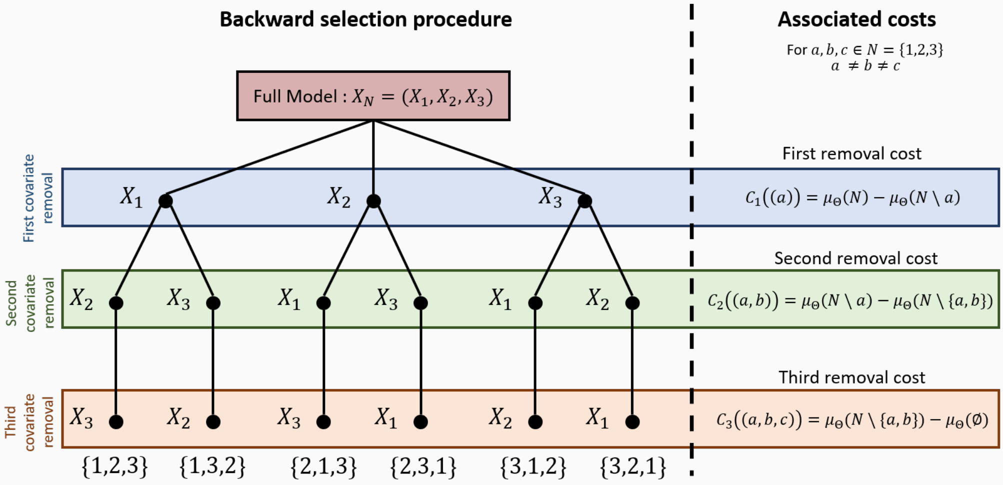
Annexe 1 : Axiome "Limit Proper Exclusion"

Axiome (Limit Proper Exclusion)

Let w be defined by a model Θ and performance measure μ where $\beta_i^* = 0$. Consider a sequence of games w_k, where $\beta_j^k = \beta_j^*$ for $j \neq i$. Assume that $\beta_i^1 > 0$ and $\beta_i^k \to 0$. Then :

$$\lim_{k \to \infty} \phi_i(w_k) = 0 \tag{25}$$

Annexe 2 : Illustration Backward



Annexe 3 : Régression logistique

In the case of logistic models, the following R^2 performance metric is computed :

$$R^2(S) = 1 - \frac{D(Y, X_S)}{D(Y, \emptyset)} \tag{26}$$

where $D(.,.)$ denotes the nested model's deviance, and $D(Y, \emptyset)$ the null model deviance (e.g., by only including an intercept). The deviance is defined by :

$$D(Y, X_S) = -2\log\left(\mathcal{L}(\widehat{\beta}_S)\right) \tag{27}$$

where $\mathcal{L}$ denotes the likelihood ($\mathcal{L}(\widehat{\beta}_S)$ denotes the value of the maximized likelihood, since $\widehat{\beta}_S$ are assumed to be MLE estimators of β_S) of the nested model, given an i.i.d. data sample.

Annexe 4 : Régression logistique 'Airquality'

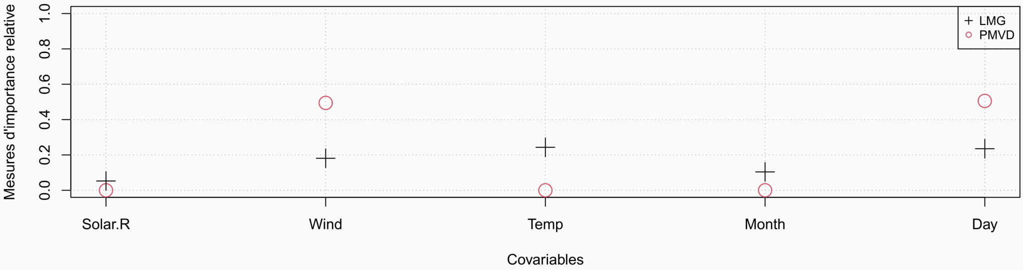
$$Y = \mathbb{1}_{\mathrm{Ozone} \geq 120}$$

Performance : $R^2(N) = 1.$

Part de $Y = 1$ **sur les données :** 2.7%

```
#Logistic Regression
LMG.airqlog <- lmg(X=airq[,-1],
                   y=airq$Ozone>=120,
                   logistic=T)
PMVD.airqlog <- emvd(X=airq[,-1],
                   y=airq$Ozone>=120,
                   logistic=T)
```

Listing 2: R code Airquality.



Annexe 5 : Identification de la PMD

- From the Limit Proper Exclusion, one has that $\frac{\partial \ln L(r)}{\partial \ln w(S)} < 0$.
- Since $\frac{\partial X}{\partial Y} = \frac{\partial X}{\partial \ln Y} \times \frac{1}{Y}$, this leads, to $-\frac{\partial \ln L(r)}{\partial w(S)} = \frac{1}{w(S)}$.

In turn, these observations lead to :

$$-\ln L(r) = c_r + \sum_{S \in r} \int_0^{w(S)} \frac{1}{x} dx = c_r + \sum_{S \in r} \ln w(S). \tag{28}$$

where c_r is a multiplicative factor dependent of r. However, the anonymity axiom requires that c_r should be constant for all $r \in \mathcal{R}(N)$, and appears both in the numerator and denominator of $p(r)$. One could subsequently assume that $c_r = 0$. This leads to the unique identification of $L(r), \forall r \in \mathcal{R}(N)$ as being :

$$L(r) = \left(\prod_{S \in r} w(S) \right)^{-1} \tag{29}$$