UNIVERSITÉ
DE MONTPELLIER

ECONOMÉTRIE SOUS PYTHON

COURS – DU PYTHON

2021 - 2022

SAMUEL STOCKSIEKER

SOMMAIRE

UNIVERSITÉ DE MONTPELLIER

- ANALYSE DE DONNÉES
- MODÉLISATION
- SERIES TEMPORELLES

UNIVERSITÉ DE MONTPELLIER

Rappel du processus standard de modélisation

1. Extraction des données / construction des bases
2. Analyse exploratoire des données
   - Analyse et traitement du format de la base de données : compréhension des données, contrôles de cohérence et de complétude, gestion des anomalies, etc.
   - Analyse descriptives des données : univariée, multivariée, Visualisation des données, analyse statistiques des données (classification, réduction de dimension, regroupement de modalités, etc.)
3. Modélisation (classification, régression, clustering): choix, sélection (recodage, regroupement, transformation, filtre variables) et validation
4. Interprétation et analyse des résultats
5. Prédiction
6. Maintenance

ANALYSE DE DONNÉES

UNIVERSITÉ DE MONTPELLIER

Analyse de format

- Dimensions : nombre de lignes, nombre de colonnes, etc.
- Typage des variables
- Nom des variables

Accès aux données

- Elément
- Variable (colonne)
- Observations (ligne)
- Filtre : unique, combiné
- Scission de données (sous–groupe)
- Tri

ANALYSE DE DONNÉES

UNIVERSITÉ DE MONTPELLIER

Analyse univariée

- Tableau de fréquence (absolu/relatif)
- Histogramme
- Statistiques descriptives, quantiles
- Contrôles des DI/DM (%dictionnaire des données)

Analyse multivariée

- Tableau de fréquence croisé
- Graphiques : courbes superposées, histogrammes, nuage de points
- Boites à moustaches parallèles
- Corrélation
- Diagnostic de cohérence croisée

ANALYSE DE DONNÉES

UNIVERSITÉ DE MONTPELLIER

Coefficients de corrélation

- Pearson : relation linéaire, paramétrique
- Spearman/Kendall : relation monotone, non paramétrique

Test d'adéquation

- Test de Student [P] : (deux échantillons non apparié/indep gaussiens de même variance ou non) ; H_0 : les échantillons ont même moyenne
- Test de Fisher [P] (deux échantillons gaussiens) ; H_0 : les échantillons ont la même variance
- Test de Wilcoxon [NP]: H_0 : les 2 distributions sont « proches »
- ANOVA [P] : (hypothèse de Normalité et indépendance), H_0 : les x échantillons ont même moyennes
- Test de Kruskal [NP] : H_0 : les échantillons ont mêmes paramètres de position

UNIVERSITÉ DE MONTPELLIER

Analyse en Composantes Principales

- Représenter graphiquement les individus sur un plan via la création de nouvelles variables définie à partir des variables initiales ⇔ construire un système de représentation de dimension réduite préservant les distances entre individus

- Déformer le moins possible la réalité : proximités entre individus préservées avec un nombre de composantes suffisamment représentatives (% d'inertie exprimée)

- Permet d'évaluer les relations entre les variables (quantitatives)

- Permet d'analyser les ressemblances entre individus

- Evite la redondance d'information (si corrélation)

UNIVERSITÉ DE MONTPELLIER

Analyse Factorielle Discriminante

- Expliquer une variable qualitative par d'autres variables (quantitatives) : proposer un nouveau système de représentation
- Vérifier la pertinence des groupes : discerner le plus possible les groupes dans le nouvel espace de représentation
- Identifier les caractéristiques des groupes à l'aide de variables explicatives
- Prédire le groupe d'appartenance d'un nouvel individu

UNIVERSITÉ DE MONTPELLIER

Analyse Factorielle des Correspondances

- Etudier la relation entre 2 variables qualitatives
- Résumer et visualiser l'information contenue dans un tableau de contingence des 2 variables
- Identifier des similitudes de profils à partir du tableau de contingence
- Visualisation graphique une association entre les profils « lignes » et les « profils colonnes »

ANALYSE DE DONNÉES

UNIVERSITÉ DE MONTPELLIER

Analyse des Correspondances Multiples

- Analyse des individus décrit par une série de variables qualitatives (« ACP » pour les variables quantitatives)
- Permet d'étudier la relation entre les variables
- Représentation graphique des variables et des observations

Classification Hiérarchique

- Regrouper (agréger) les individus au sein de classes : homogénéité intra-classes et hétérogénéité interclasses
- Donne une mesure de la proximités entre les sous-groupes

SOMMAIRE

UNIVERSITÉ DE MONTPELLIER

- ANALYSE DE DONNÉES
- MODÉLISATION
- SERIES TEMPORELLES

PRIME PURE EN NON VIE

UNIVERSITÉ DE MONTPELLIER

Modèle collectif

La prime pure, défini au niveau de la police, est la charge attendu par police, soit :

$$\pi = \frac{E\left(S^{coll}\right)}{n} = \frac{E(E[S|N])}{n} = \frac{E\left(N.E(X)\right)}{n} = E\left(\frac{N}{n}\right).E(X)$$

$E\left(\frac{N}{n}\right)$ représente la fréquence (nombre de sinistre moyen) par police

$E(X)$ représente le coût moyen d'un sinistre

Dans le cadre du modèle collectif, l'espérance de la charge de sinistre par police i.e. la prime pure, peut être obtenue par le produit des espérances de la fréquence et du coût : modèle « fréquence – coût »

Obtenu facilement, sans simulations numériques, ce résultat est un pilier de la tarification en assurance Non-Vie. Il permet de calculer des primes pures mais est insuffisant pour calibrer finement la distribution de S nécessaire aux calculs de probabilité de ruine ou en réassurance

PRIME PURE EN NON VIE

UNIVERSITÉ DE MONTPELLIER

Retour sur les hypothèses du modèle collectif

- $\{X_i$ v.a.i.i.d$\}$: *le montant individuel des sinistres ont le même comportement aléatoire ; Le montant individuel d'un sinistre n'influe pas sur les autres montants individuels de sinistre*
  → N'est possible que si le facteur d'actualisation peut être négligé et s'il n'y a pas de risque de dérive du coût des sinistres. Or, avec l'inflation, les coûts sont d'autant plus impactés que la durée est longue : ils ne seront pas stationnaires sauf actualisation des montants au taux adéquat. L'indépendance des sinistres peut s'avérer impossible dans certains cas (sinistres communs). Cette hypothèse nécessite donc de segmenter le portefeuille…
- $\{X_i\}$ indépendants de N : *le nombre de sinistres n'a pas d'incidence sur les montants individuels des sinistres*
  → Valable si le portefeuille est homogène, ce qui n'est pas le cas dans certaines raisons (*e.g* géographique)

PRIME PURE EN NON VIE

UNIVERSITÉ DE MONTPELLIER

Tarification en assurance Non Vie

Pour déterminer la prime pure $E(S)$, il suffit donc de déterminer la fréquence moyenne $E(N)$ et le coût moyen $E(X)$

Nombre de sinistres : N

- Loi Poisson : moyenne = variance (+ lois poisson-mélange)
- Loi binomiale : variance < espérance
- Loi binomiale négative : variance > espérance
- ZIP, ZINB
- Quasi-poisson, quasi-binomiale

Coût des sinistres : X

- Loi Normale, inverse-Gaussienne
- Loi Log-Normale
- Loi Gamma
- Loi Pareto (réassurance)
- Etc.

MODÉLISATION

UNIVERSITÉ DE MONTPELLIER

Processus classique de modélisation

1. Construction de la base de données : extraction, fusion, etc.
2. Analyse et traitement du format de la base de données : compréhension des données, contrôles de cohérence et de complétude, gestion des anomalies, etc.
3. Analyse descriptives des données : univariée, multivariée, analyse de données (classification, réduction de dimension, etc.)
4. Régression
    1. choix du modèle
    2. Sélection du modèle
    3. Validation du modèle : analyse de l'adéquation / significativité
5. Prévision - avec lissage éventuel
6. Mise à jour du modèle

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Objectif d'un modèle de régression

Expliquer une variable endogène Y par p variables exogènes $X_j, j = 1, \dots, p$

Ecriture du modèle de régression linéaire multiple

$$Y = X\alpha + \epsilon$$

$$y_i = \alpha_0 + \sum_{j=1}^{p} \alpha_j . x_{i,j} + \epsilon_i$$

- Il faut estimer les $(p + 1)$ paramètres $(\alpha_0, \dots, \alpha_p)$ à partir d'un échantillon de n observations
- ϵ représente l'erreur du modèle : toutes les insuffisances du modèle : écart valeurs prédites et valeurs observées
- Le modèle est dit additif

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Processus de construction du modèle

1. Estimation des paramètres (α) sur la base de l'échantillon
2. Mesure du pouvoir explicatif du modèle (globalité) : tableau d'analyse de la variance, R^2
3. Test de significativité globale du modèle
4. Test de significativité des coefficients : apport marginal des variables explicatives
5. Sélection des variables
6. Validation du modèle
7. Evaluation du pouvoir prédictif du modèle
8. Interprétation des résultats (coefficients)
9. Prévision : application du modèle

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Hypothèses du modèle

Hypothèses (stochastiques) du modèle :

- X_j non aléatoires
- $E(\epsilon_i) = 0$: espérance de l'erreur est nulle ≈ le modèle est bien spécifié en moyenne
- $E(\epsilon_i^2) = \sigma_\epsilon^2$: la variance de l'erreur est constante (homoscédasticité) et $\mathrm{COV}(\epsilon_i, \epsilon_{i'}) = 0$ pour $i \neq i'$ i.e. erreurs indépendantes : non – autocorrélation des résidus
- $COV(x_{ij}, \epsilon_i) = 0$ l'erreur est indépendante des variables exogènes
- $\epsilon_i \sim N(0, \sigma_\epsilon)$: les erreurs sont distribuées selon une loi Normale

Hypothèses structurelles du modèle : (XX') régulière ⇔absence de colinéarité entre les exogènes, $n > p + 1$, etc.

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Tableau d'analyse de variance et coef. de détermination

Décomposition de la variabilité de Y :

- $SCE = \sum_i (\hat{y}_i - \bar{y})^2 \equiv$ variabilité expliquée par le modèle
- $SCR = \sum_i (y_i - \hat{y}_i)^2 \equiv$ variabilité résiduelle
- $SCT = \sum_i (y_i - \bar{y})^2 \equiv$ variabilité totale

Coefficient de détermination :

$$R^2 = \frac{SCE}{SCT}$$

- $0 \leq R^2 \leq 1$: plus il tend vers 1, meilleur est le modèle
- Augmente avec le nombre de variables explicatives, même non pertinentes
- On peut alors utiliser le R^2 corrigé (pénalisation par les ddl) mais on préféra l'AIC et le BIC (permet de détecter une colinéarité)

$$AIC = nln\left(\frac{SCR}{n}\right) + 2(p+1) \; ; BIC = nln\left(\frac{SCR}{n}\right) + \ln(n)(p+1)$$

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Test de significativité globale du modèle

- Formulation : on veut tester si le modèle, dans sa globalité, est pertinent. L'hypothèse nulle est qu'aucune des exogènes n'apporte de l'information pour expliquer Y donc :

$$\begin{cases} H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0 \\ H_1 : \exists j \quad t.q \quad \alpha_j \neq 0 \end{cases}$$

- Statistique de test :

$$F = \frac{R^2/p}{(1-R^2)/(n-p-1)} \sim F(p, n-p-1) \; sous \; H_0$$

- Ce principe de test peut également servir à comparer 2 modèles emboités : test de significativité d'un bloc de coefficients

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Test de significativité d'un coefficient

- Formulation : on souhaite évaluer la pertinence des variables, indépendamment des autres (impact marginal). En supposant que $\epsilon_i \sim N(0, \sigma_\epsilon)$, on a :

$$\begin{cases} H_0 : \alpha_j = 0 \\ H_1 : \alpha_j \neq 0 \end{cases}$$

- Statistique de test :

$$\frac{\widehat{\alpha_j} - \alpha_j}{\widehat{\sigma_{\widehat{\alpha_j}}}} = \frac{\widehat{\alpha_j}}{\widehat{\sigma_{\widehat{\alpha_j}}}} \sim T(n - p - 1) \; sous \; H_0$$

- Sous ces hypothèses, il devient possible de déterminer un IC du coefficient : $\alpha_j \pm t_{1-\alpha/2} \times \widehat{\sigma_{\widehat{\alpha_j}}}$

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Analyse des résidus

- Résidus : estimation du terme d'erreur ; i.i.d $\sim N(0, \sigma_\epsilon)$
- Retour sur les hypothèses : identifier toutes forme de « régularité » dans les résidus et/ou dépendance avec les variables
- Une premier aperçu graphique des « y observés » vs « y prédits » permet d'indiquer la pertinence du modèle (travailler sur une base test et non d'apprentissage)
- Graphique « endogène vs résidus » : corrélation? Dispersion ? Points atypiques/influents ? Homogénéité ? Equirépartition autour de 0 vs symétrie ?
- Résidu s'écartant significativement ≡ observation atypique/mal modélisée (détection des points influent assez complexe)
- Graphique « exogènes vs résidus » : points atypiques ? Dépendance ? Dispersion selon les valeurs de X ? Homoscédasticité ?
- [Données longitudinales : graphique « temps vs résidus »]

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Analyse des résidus

- Non –linéarité des résidus observée : relation non linéaire, transformer une ou plusieurs variables exogènes
- Rupture de structure (résidus en « blocs » : ensemble de définition distincts/points d'inflexion) ➔ liens « endo - exo» non constants – plusieurs régression à mener ? Test de Chow, tests de stabilité, etc.
- Hétéroscédasticité : variance des résidus non constante et dépendante d'une exogène (∃ tests stats)
- Autocorrélation : spécifique aux données longitudinales (ST)
- Tests d'adéquation à la loi Normale :
  - QQ-plot (quantiles emp. vs quantiles th. : droite ⇔comptabilité entre les distributions)
  - Test de jarque-Bera : basé sur les coef. d'asymétrie et applatissement $\left(\gamma_1 = \frac{\mu^3}{\sigma^3}\,; \gamma_2 = \frac{\mu^4}{\sigma_4} - 3\right) : H_0 : \epsilon \sim N \Rightarrow \gamma_1 = \gamma_2 = 0$
  - Test de SW : H_0 : la distribution de l'échantillon est Gaussienne

RAPPEL : REG. LIN. MUL.

UNIVERSITÉ DE MONTPELLIER

Prédiction du modèle

- Valeur prédite :

$$\widehat{y_{i^*}} = \widehat{\alpha_0} + \sum_j^p \widehat{\alpha_j} \times x_{i^*,j}$$

- Estimation de la variance de l'erreur de prédiction :

$$\frac{\widehat{\epsilon_{i^*}}}{\widehat{\sigma_{\epsilon_{i^*}}}} = \frac{\widehat{y_{i^*}} - y_{i^*}}{\widehat{\sigma_{\epsilon_{i^*}}}} \sim T(n-p-1)$$

$$\Rightarrow \widehat{y_{i^*}} \pm t_{1-\alpha/2} \times \widehat{\sigma_{\epsilon_{i^*}}}$$

- La variance (donc l'IC) sera d'autant plus grand sur la régression est de mauvaise qualité

RAPPEL : GLM

UNIVERSITÉ DE MONTPELLIER

Ecriture du modèle

$$g(y_i) = g\big(E(Y|X_i)\big) = \alpha_0 + \sum_{j=1}^{p} \alpha_j . x_{i,j} + \epsilon_i$$

- $g()$ est la fonction de lien
- Y doit être une loi de la famille exponentielle
- $\alpha_0 + \sum_{j=1}^{p} \alpha_j . x_{i,j}$ est le prédicateur linéaire
- ϵ_i est le terme d'erreur

RAPPEL : GLM

UNIVERSITÉ DE MONTPELLIER

Significativité globale du modèle

- Déviance : écart entre la log-vraisemblance obtenue avec le modèle et celle obtenue avec un modèle parfait dit « saturé »

$$D = 2\phi \times [lnL(y) - \ln L(\hat{\mu})] \quad ; \quad \hat{\mu} = g^{-1}(X\alpha)$$

➔ Théoriquement, plus cette quantité est faible, meilleur est le modèle

- Déviance normalisée (*scaled deviance*) : déviance normalisée par le paramètre de dispersion ϕ

$$D = 2 \times [lnL(y) - \ln L(\hat{\mu})]$$

- Statistique du χ^2 de Pearson

$\chi^2 = \sum \frac{(y_i - \widehat{\mu_i})^2}{V(\widehat{\mu_i})} \sim \chi^2(n-p)$ sous H_0 : le modèle est bien spécifié

SOMMAIRE

UNIVERSITÉ DE MONTPELLIER

- ANALYSE DE DONNÉES
- MODÉLISATION
- **SERIES TEMPORELLES**

UNIVERSITÉ DE MONTPELLIER

Généralités

- Série temporelle / série chronologique : processus aléatoire dans le temps (S_t) : suite finie de données indexées dans le temps
- Objectif : modéliser le processus afin de l'expliquer et le prédire
- Approche classique : Tendance et saisonnalité : $S_t = m_t + s_t + X_t$ (mod. additif), m_t représente la tendance : fonction déterministe représentant la variation long terme (le plus lisse possible),
  s_t représente la saisonnalité, composante périodique, (fonction déterministe) de période p tq $\sum_{i=1}^{p} s_{t+1} = 0$
  X_t représente le bruit (résidus) aléatoire que l'on espère stationnaire
- Une transformation de la série est parfois nécessaire pour permettre sa caractérisation
- Dépendance : autocorrélation – généralement utilisé pour modéliser les résidus X_t

UNIVERSITÉ DE MONTPELLIER

Approche générale

- Analyse graphique : tendance, composante saisonnière, rupture, valeurs aberrantes, etc.
- Estimer et retirer la tendance et la saisonnalité pour isoler les résidus
  - Transformation des données éventuelle
  - Différenciation éventuelle : différence terme à terme sur un pas défini
- Modélisation du bruit (sous forme de série stationnaire, faible, i.e. espérance et covariance stables dans le temps)
- Prévision, analyse, description, explication

UNIVERSITÉ DE MONTPELLIER

Estimation de la tendance

- On pose $m_t = \sum_i^n \alpha_i f(i,t)$: tendance ≡combinaison linéaire de fonction temporelles connues et déterministes
  La tendance peut alors être linéaire, quadratique, polynomiale, exponentielle, etc.

➔ Objectif : estimer les paramètres α :

- Méthode des moindre carrés : $argmin\ (\sum_t (S_t - m_t)^2)$
- Différenciation (élimination uniquement) : $S_t - S_{t-T}$
- Moyennes mobiles (linéarité /morceau) : $m_t = \frac{1}{2q+1} \sum_{k=-q}^{q} S_{t+k}$
- Etc.

UNIVERSITÉ DE MONTPELLIER

Estimation de la saisonnalité

- On peut poser : $S_t = \sum_{j=1}^{p} \beta_j s_t^j$
- On peut fixer $S_j^t = \begin{cases} 1 \; si \; la \; période \; \text{à} \; t \; est \; j \\ 0 \; sinon \end{cases}$

➔ Objectif : estimer les paramètres β :

- Méthode des moindre carrés
- Moyennes mobiles (élimination uniquement)
- Différenciation (élimination uniquement)
- Séries de Fourrier
- Méthodes à noyau, approches polynomiales
- Splines (fonction définie par morceaux par des polynômes)
- Etc.

UNIVERSITÉ DE MONTPELLIER

Modélisation des résidus

- Fonction d'autocorrélation (AC) : $\rho_S(h) = \frac{Cov(S_t, S_{t+h})}{\sqrt{Var(S_t)Var(S_{t+h})}}$; mesure le degré de dépendance entre les valeurs de la série à différents instants
- Le tracé de la fonction $\rho_S(h)\ \forall\ h \in N$ s'appelle l'autocorrélogramme simple.
- Fonction d'autocorrélation partielle (ACP) : $\phi_S(h) = corr\big(S_h - P(S_h|S_{h-1}, \dots, S_1), S_0 - P(S_0|S_{h-1}, \dots, S_1)\big)$ avec $P(S_0|S_{h-1}, \dots, S_1) = \alpha_1 S_1 + \dots + \alpha_{h-1} S_{h-1}$ RL ; mesure la corrélation entre S_0 et S_h en ayant retiré l'explication des variables intermédiaires i.e. leur dépendance pouvant influer l'autocorrélation
- Le tracé de la fonction $\phi_S(h)\ \forall\ h \in N$ s'appelle l'autocorrélogramme partiel.

UNIVERSITÉ DE MONTPELLIER

Modélisation des résidus

- Modèle Auto Régressive
  (S_t) est un processus $AR(p)$, centré, s'il peut s'écrire sous la forme : $S_t = \epsilon_t + \sum_{j=1}^{p} \alpha_j S_{t-j}$ avec ϵ_t est un bruit blanc centré ; la valeur de S_t à t dépend d'un choc aléatoire à l'instant t, ϵ_t, indépendant de l'historique et d'une fonction linéaire de son passé // prédiction de S_t à partir des p dernières observations passées

  ➔ L'AC décroit avec h vers 0; l'ACP est nul $\forall h > \mathrm{p}$ et vaut α_p à l'ordre p

UNIVERSITÉ DE MONTPELLIER

Modélisation des résidus

- Modèle Moving Average
  (S_t) est un processus $MA(q)$ s'il peut s'écrire de la forme :
  $S_t = \epsilon_t + \beta_1 \epsilon_{t-1} + \cdots + \beta_q \epsilon_{t-q}$ avec $\epsilon_j, t-q \leq j \leq t$ des bruits blancs centrés ; la valeur de S_t à t dépend uniquement de bruits blancs

  ➔ L'AC est nul $\forall h > \mathrm{q}$; l'ACP décroit vers 0

UNIVERSITÉ DE MONTPELLIER

Modélisation des résidus

- Modèle ARMA
  (S_t) est un processus $\mathrm{AR}MA(p,q)$ s'il peut s'écrire de la forme :
  $S_t = \sum_{k=1}^{p} \alpha_k \, S_{t-j} + \sum_{j=0}^{q} \beta_j \, \epsilon_{t-j}$ avec $\beta_0 = 1$

  ➔ L'AC décroit exponentiellement avec h après l'ordre q (conditions).L'ACF décroit vers 0 (pas de caractérisation particulière)

- Modèle AR Integrated MA (ARIMA)
  Généralisation de l'ARMA pour des processus non stationnaires
  (S_t) est un processus $ARIMA(p,d,q)$ si le processus $Y_t := \Delta^d S_t$ est un processus $ARMA(p,q)$; Δ étant l'opérateur de différenciation tel que : $\Delta : S_t \to S_t - S_{t-1}$

UNIVERSITÉ DE MONTPELLIER

Modélisation des résidus

- Modèle Seasonal ARIMA (SARIMA)
  (S_t) est un processus $SARIMA(p, d, q, T)$ si le processus $Y_t := \Delta^d . \Delta_T S_t$ est un processus $ARMA(p, q)$; Δ_T étant l'opérateur de différenciation de période T tel que : $\Delta_T : S_t \rightarrow S_t - S_{t-T}$

- Modèle SARIMAX
  Modèle SARIMA avec intégration de variables explicatives (régression)

BIBLIOGRAPHIE

UNIVERSITÉ DE MONTPELLIER

Quelques sources à exploiter

- http://eric.univ-lyon2.fr/~ricco/cours/supports_data_mining.html
- https://freakonometrics.hypotheses.org/tag/notes-de-cours
- https://www.ceremade.dauphine.fr/~idris/Intro-actuariatM1.pdf
- http://www.ressources-actuarielles.net/EXT/ISFA/fp-isfa.nsf/34a14c286dfb0903c1256ffd00502d73/d084f3c15c3ea6fbc1256f15001f03fb/$FILE/ModeleCollectif.pdf
- http://www.ressources-actuarielles.net/EXT/ISFA/fp-isfa.nsf/34a14c286dfb0903c1256ffd00502d73/6c3ddda0fcbc2b5dc1256f7800690677/$FILE/Modele_collectif.pdf
- http://math.univ-lille1.fr/~suquet/Polys/AtelierJA07.pdf
- https://cel.archives-ouvertes.fr/cel-00550583/document