Statistique de la distribution des galaxies et analyse fractale

Antoine Labatie (SAp, CEA Saclay) 22/01/2010

Contexte général

- Relevés de galaxies de plus en plus grands
 - meilleure estimation des propriétés statistiques de la distribution des galaxies
- Intérêts:
 - Validation des modèles par comparaison des statistiques mesurées avec simulations
 - Contraintes sur paramètres cosmologiques (échelle d'homogénéité, prépondérance de certaines structures aux ≠ échelles)



Modèle statistique pour la distribution des galaxies

- Processus ponctuel: X=(x₁,...,x_N) où N est une variable aléatoire et les x_i sont aussi aléatoires (positions des galaxies)
- Modèle statistique \rightarrow processus de Cox en 2 étapes:
 - Densité de matière noire sous-jacent
 ρ(x) réalisation d'un champ aléatoire (hypothèse de champ *homogène, isotrope*)
 - L'apparition de galaxies suit un processus de Poisson avec une intensité locale λ(x) qui dépend du champ de densité ρ(x)



Fonction de corrélation à 2 points

• Proba de présence d'une galaxie dans δV :

$$\delta P = \bar{n} \delta V$$

• Si on connaît l'emplacement d'une galaxie, la probabilité de présence d'une autre galaxie dans volume δV à une distance r est:

$$(\delta P)_p = \bar{n}(1 + \xi(r))\delta V$$



 $\delta P_{12} = \bar{n}^2 (1 + \xi(r_{12})) \delta V_1 \delta V_2$

 $\xi(r)$ donne la tendance au regroupement des points (statistique d'ordre 2)



Fonction de corrélation à 2 points

• Si on suppose que le champ de densité donne exactement la densité de Poisson de la distribution, i.e. $\lambda(\mathbf{x}) = \rho(\mathbf{x})$:

$$\delta P_{12} = \langle \rho(\mathbf{r_1}) \delta V_1 \rho(\mathbf{r_2}) \delta V_2 \rangle$$
$$\delta P_{12} = \bar{\rho}^2 (1 + \xi(r_{12})) \delta V_1 \delta V_2$$

• Alors:

$$\xi(r_{12}) = \frac{1}{\bar{\rho}^2} \operatorname{Cov}(\rho(\mathbf{r_1}), \rho(\mathbf{r_2}))$$

• En notant les fluctuations $\delta(r)$:

$$\delta(\mathbf{r}) = \frac{\rho(\mathbf{r}) - \bar{\rho}}{\bar{\rho}} \qquad \qquad \xi(r_{12}) = \operatorname{Cov}(\delta(\mathbf{r_1}), \delta(\mathbf{r_2}))$$

Fonction de corrélation à 2 points: problèmes dans l'estimation

- Si on change seulement la densité moyenne du champ, il n'y a pas de changement de la fonction de corrélation
- Problème:
 - Les quantités auxquelles on a facilement accès mélangent les contributions de \bar{n} et $\xi(\mathbf{r})$:

$$(\delta P)_p(r) = \bar{n}(1+\xi(r))\delta V$$

 On peut estimer chaque quantité, puis faire la division, mais l'estimateur est alors biaisé

Spectre de puissance

• Spectre de puissance défini comme la puissance moyenne pour un mode de Fourier:

$$\delta(\mathbf{r}) = \frac{\rho(\mathbf{r}) - \bar{\rho}}{\bar{\rho}} , \quad \hat{\delta}(\mathbf{k}) = TF(\delta)$$
$$P(\mathbf{k}) = \langle |\hat{\delta}(\mathbf{k})|^2 \rangle$$

• C'est la transformée de Fourier de la fonction de corrélation:

 $P(\mathbf{k}) = TF(\xi)$

- Comme pour la fonction de corrélation, l'isotropie du champ entraine l'isotropie du spectre de puissance
- Pour un champ gaussien, les modes de Fourier sont statistiquement orthogonaux (i.e. indépendants dans ce cas gaussien):

$$\langle \hat{\delta}(\mathbf{k})\hat{\delta}(\mathbf{k}') \rangle = \delta(\mathbf{k} - \mathbf{k}')$$

Estimation du spectre de puissance

• On observe une réalisation du champ sur une fenêtre (géometrie pouvant être complexe)



• En Fourier convolution des modes avec une fonction correspondant a la fenêtre:

$$\mathbb{E}[\hat{P}(\mathbf{k})] = \int \frac{d^3 \mathbf{k}'}{2\pi^3} P(\mathbf{k}') |G(\mathbf{k} - \mathbf{k}')|^2$$

 Pour estimer le spectre de puissance on se sert de l'isotropie → on moyenne l'énergie des k sur une enveloppe sphérique:

$$\hat{P}(k) = \frac{1}{V_k} \int_{V_k} d^3 \mathbf{k}' P(\mathbf{k}')$$

• La convolution des modes peut poser des problèmes pour l'estimation.

Possible méthode d'estimation du spectre de puissance

- Transformée en ondelettes du champ de densité
- Les bandes du spectre de puissance sont reliées à la variance des coefficients d'ondelettes

$$P(\Delta k_i) = \int_{\Delta k_i} P(k) |\hat{\psi}_i(k)|^2 = \operatorname{Var}(w_i(\mathbf{x}))$$

- On peut alors théoriquement retrouver la fonction de corrélation par TF⁻¹
 - algorithme rapide en N log(N) alors que méthodes classiques sont en N²
 - mais incertitude sur la précision des résultats et sur le biais ?

Propriétés fractales de la distribution

• Pour une distribution fractale de dimension $d_F < d$ (dimension de l'espace d=3), la masse contenue dans les boules de tailles R est:

$$M(R) \propto R^{d_F}$$

• Pour des boules centrées sur les galaxies:

$$M(R) = \bar{n} \int_0^R 4\pi s^2 (1 + \xi(s)) ds$$
$$\xi(r) \propto r^{d_F - d}$$

- On observe une telle loi de puissance avec $d_F \approx 2$, mais seulement à petite échelle
- À plus grande échelle, on observe un plateau





Controverse sur l'homogeneité de l'Univers

• Un groupe de chercheurs défend que l'Univers est une fractale à toutes les échelles

---> cela ne se verrait pas sur la fonction de correlation à grande échelle à cause des effets de volume fini

- Argumentation sur l'échelle de corrélation r_0 : $\xi(r_0) = 1$
 - Pour une vraie fractale on ne peut pas définir $ar{n}$
 - L'estimation de la fonction de corrélation utilisant le \bar{n} de l'échantillon doit présenter une longueur de corrélation qui varie avec la taille de l'échantillon R: $r_0 \propto R$

$T_0 \propto$

- Cependant:
 - Cette relation ne semble plus s'appliquer aux relevés plus grands actuels
 - Autres arguments forts pour l'homogénéité: isotropie du fond diffus cosmologique, isotropie de certains types de source, etc

Analyse multifractale

• Même si la distribution n'est pas une pure fractale, elle peut avoir une nature multifractale à certaines échelles

$$|f(\mathbf{x}) - P_n(\mathbf{x} - \mathbf{x}_0)| \le C|\mathbf{x} - \mathbf{x}_0|^{h(\mathbf{x}_0)}$$

$$D(h) = d_H \{ \mathbf{x} \in \mathbb{R}^2, \ h(\mathbf{x}) = h \}$$

- Différentes méthodes pour déterminer le spectre de singularités:
 - "box counting"
 - "Maxima des maxima du module de la transformée en ondelettes" (MMMTO):

$$\mathcal{Z}(q,a) = \sum_{\mathcal{L} \in \mathcal{L}(a)} \left(\mathcal{M}_{\psi}[f](\mathbf{r},a) \right)^q \qquad \mathcal{Z}(q,a) \sim a^{\tau(q)} , \ a \to 0^+$$
$$D(h) = \min_q \left(qh - \tau(q) \right)$$

- Mais:
 - difficultés algorithmiques
 - difficultés dans le cas présent pour interpréter le spectre (seule possibilité serait de comparer le spectre obtenu avec les spectres de simulations cosmologiques)

Modification de la méthode pour l'analyse a différentes échelles

• Le module de la TO a une variation reliée a l'exposant de Hölder:

$$\mathcal{M}_{\psi}[f] (\mathcal{L}_{\mathbf{x}_0}(a)) \sim a^{h(\mathbf{x}_0)}, \quad a \to 0^+$$

• Une singularité donne une ligne de maxima à chaque échelle:

$$\mathcal{M}_{\psi}[f](\mathcal{L}_{\mathbf{x}_1}(a)) \sim a^{n_{\psi}}, \quad a \to 0^+$$

- Extension à différentes échelles:
 - on devrait pouvoir trouver la dimension fractale locale à chaque echelle (déterminer les types de structures, i.e. amas, filaments, feuillets, ainsi que leurs proportions)
 - on devrait pouvoir tester s'il y a homogénisation en regardant s'il y a une rupture dans l'exposant de variation des coefficients en ondelettes ($d_F=3$)

