Les plus grandes valeurs propres de matrices aléatoires

S. Péché ,
Université Paris 7,

10/10/2011,
Séminaire "Point de vue",
Université Paris 7

Plan

Première partie

- Matrices aléatoires: définitions, premières propriétés.
- Comportement "standard" du bord du spectre.
- Modèles déformés , à queue lourde (universalité ou non).

Seconde partie

- Application à des problèmes d'estimation de covariance
- Rencontres avec la loi de Tracy-Widom: classe d'universalité (Processus déterminantaux (test))

Modèles

Soit μ (resp μ') une loi centrée sur $\mathbb{C}$ (resp. sur $\mathbb{R}$) de variance finie.

- matrices aléatoires "de Wigner": hermitiennes (ou réelles symétriques)

$$H_N = \frac{1}{\sqrt{N}}(H_{ij})_{i,j=1}^N, \quad H = H^*,$$

avec $H_{ij}, 1 \leq i < j \leq N$ variables aléatoires i.i.d. complexes (ou réelles) de loi μ. $H_{ii}, 1 \leq i \leq N$: variables aléatoires i.i.d. réelles $\perp\!\!\!\perp$ $H_{ij}, 1 \leq i < j \leq N$, de loi μ'.

Modèle né en physique avec Wigner: modélisation des niveaux d'énergie des atomes lourds.
Spectre en grande dimension: $N \to \infty$?

Modèles 2

- Matrices de covariance empirique blanches:

$$M_N = \frac{1}{N} X X^*, \ X = (X_{ij}), \ i = 1, \ldots, N, \ j = 1, \ldots, p, \ \text{avec } p = p(N),$$

  avec X_{ij} i.i.d. de loi μ.

- Matrices non blanches: $\Sigma > 0$ $N \times N$ réelle symétrique (déterministe ou non)

$$M_N(\Sigma) = \frac{1}{N} \Sigma^{1/2} X X^* \Sigma^{1/2},$$

  avec X comme au dessus.

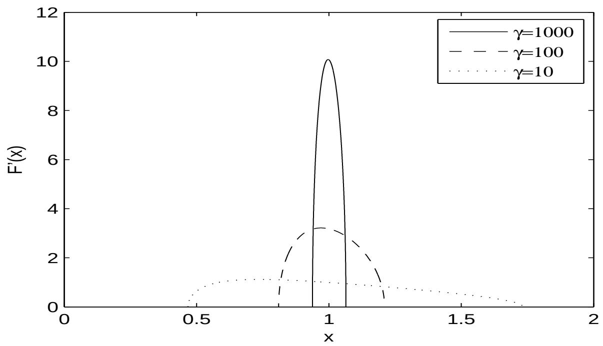
Spectre en grande dimension: $N \to \infty$ et $\exists\ \gamma \in [0, \infty]$ $\lim_{N \to \infty} p/N = \gamma$.

Motivations I

En statistique Un des problèmes importants: connaissant $M_N(\Sigma)$ quelles informations peut-on en déduire sur Σ?

-si N fixé et $p \to \infty$: $M_N(\Sigma)$ bon estimateur de Σ;

-en grande dimension (génétique, finance, grandes capacités de stockage), comportement de l'ACP?



Densité des valeurs propres de $M_N(\Sigma)$ quand $\Sigma = Id$.

Dispersion des valeurs propres: $M_N(\Sigma)$ mauvais estimateur de Σ

Motivations II.

En théorie de la communication "CDMA" : signal reçu $r = \sum_{k=1}^{p} b_k s_k + w,$
avec p le nombre d'utilisateurs,
$b_k \in \mathbb{C}$, $\mathbb{E}b_k = 0$, $\mathbb{E}|b_k|^2 = p_k$ le signal transmis,
$s_k \in \mathbb{C}^N$ la signature aléatoire
et $w \in \mathbb{C}^N$ un bruit gaussien dont les composantes sont i.i.d. $\mathcal{N}(0, \sigma^2)$.

Il s'agit de décoder le signal b_k. Une mesure de la performance du canal de communication est le "SIR" (Signal to Interference Ratio): recepteur linéaire $\hat{x}_1 = c_1^* r$

$$SIR = \frac{|c_1^* s_1|^2 p_1}{|c_1|^2 \sigma^2 + \sum_{i \geq 2} |c_1^* s_i|^2 p_i}.$$

$\Longrightarrow$ quand $N, p \to \infty$, $p/N \to \gamma$, le SIR est une fonction des valeurs propres

et des vecteurs propres de SDS^* où $S = [s_2, \ldots, s_p]$ est la matrice des signatures (aléatoires) $D = \mathrm{diag}(p_2, \ldots, p_N)$.

Motivations III.

En finance Un portefeuille $\mathcal{P}$ de N actions avec des poids w_i, $i = 1, \dots, N$.
Σ la matrice de covariance des rendements.

$$\text{Variance quotidienne à minimiser : } R^2 = \sum_{i,j=1}^{N} w_i w_j \Sigma_{ij}.$$

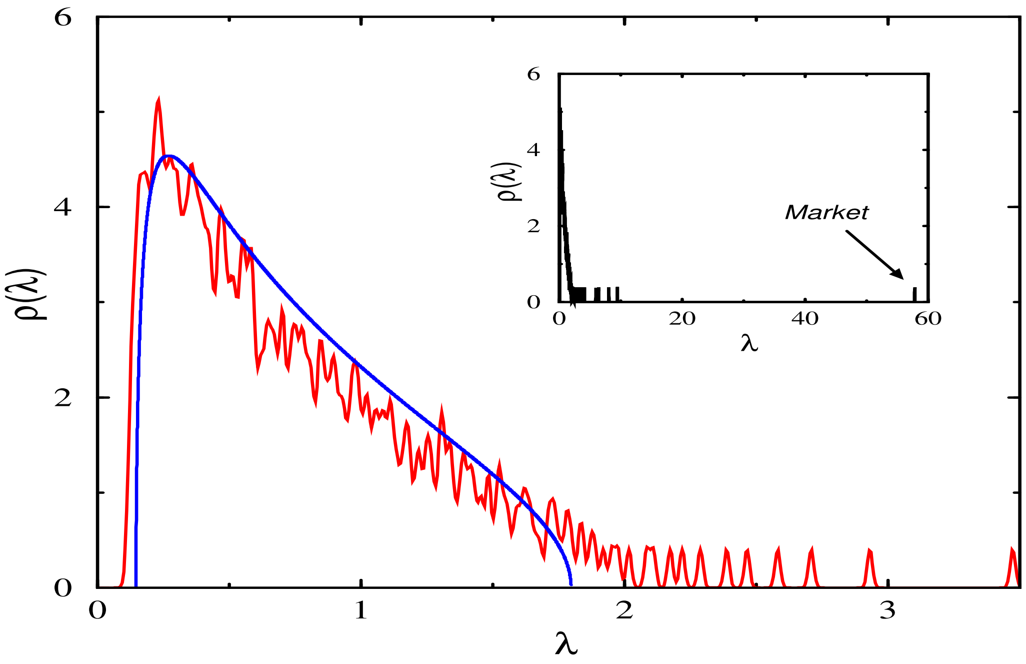
$\mathcal{P}$ optimal: poids maximal sur les vecteurs propres associés aux plus petites valeurs propres de Σ.

Problèmes : Étant donnée Y, la matrice $N \times p$ des rendements observés sur une période de p jours et la matrice de covariance empirique YY^*,

1. estimer Σ (test: $\Sigma = Id$ ou identifier des secteurs),
2. identifier les vecteurs propres associés aux plus grandes et plus petites valeurs propres.

Exemple

Bouchaud-Potters-Laloux Financial Applications of RMT (2005)



Distribution empirique des v.p pour 406 fonds du S&P 500, et comparaison avec une loi de MP. Noter la présence d'une très grande v.p. (et de qques autres).

Comportement global du spectre

Matrices de covariance empirique Soient $\pi_1 \geq \pi_2 \geq \cdots \geq \pi_N$ v.p. de Σ.

Hyp: $\rho_N(\Sigma) := \frac{1}{N}\sum_{i=1}^{N} \delta_{\pi_i} \overset{p.s.}{\to} H$ avec H une mesure de probabilité.

Soit $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ les valeurs propres de $M_N(\Sigma)$; $\mu_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}$.

Théorème Marchenko-Pastur (67)
Supposons $\sigma^2 := \mathrm{Var}(\mu) = 1$. P.s. $\lim_{N\to\infty} \mu_N = \rho_{MP}$, la loi ρ_{MP} caractérisée par sa transformée de Stieltjes:

$$z \in \mathbb{C}, \Im(z) > 0, \quad m_\rho(z) := \int \frac{1}{\lambda - z} d\rho_{MP}(\lambda);$$

Alors $m_\rho(z) = \int_{-\infty}^{+\infty} \{\tau\,[\gamma - 1 - z\, m_\rho(z)] - z\}^{-1}\, dH(\tau)$.

Valable pour les matrices réelles et complexes; Universalité (unique hypothèse sur la variance).

Comportement global du spectre

Dans le cas où $H = \delta_1$ ($\Sigma = Id$): expression explicite de la densité.
Si Var$(\mu) = \sigma^2$ et $\gamma \geq 1$,

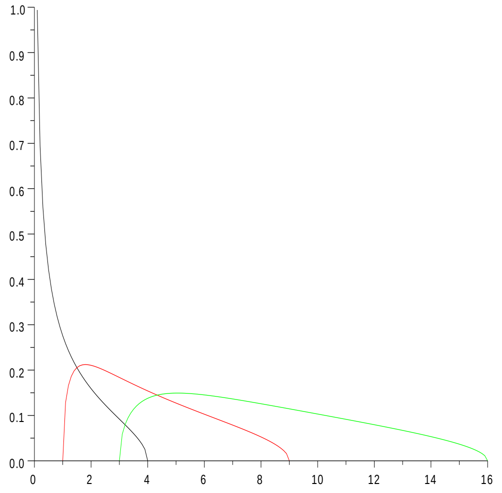


Figure 1: Densité de la loi de MP pour $\gamma = 1, 2, 4$

$$\frac{d\rho_{MP}}{du} = \frac{1}{2\pi u \sigma^2}\sqrt{(u_+ - u)(u - u_-)},$$

avec $u_\pm = \sigma^2(1 \pm \sqrt{\gamma})^2$.

Si $\gamma \leq 1$, masse en 0 de poids $1 - \gamma$.

:: Définition

Outils de la preuve:

- approche par la résolvente/ transformée de Stieltjes:

$$\frac{1}{N}\mathrm{Tr}(M_N - zI)^{-1} = \int \frac{1}{\lambda - z} d\mu_N(\lambda).$$

$M_N = 1/N \sum_k C_k C_k^*$, C_k $k$ième vecteur colonne de $\Sigma^{1/2}X$ et identité de la résolvente.

- approche par les moments (plutôt si $\Sigma = Id$)

$$\int x^k d\mu_N = \frac{1}{N}\mathrm{Tr} M_N^k = \frac{1}{N^{k+1}} \sum_{i_0,\ldots,i_{k-1},j_1,\ldots,j_k} X_{i_0 j_1}\overline{X_{i_1 j_1}} \cdots X_{i_{k-1} j_k}\overline{X_{i_0 j_k}}.$$

Montre $\mathbb{E}\int x^k d\mu_N \to \int x^k d\rho_{MP}$ et Var$\int x^k d\mu_N = O(N^{-1-\epsilon}))$.

Questions

- Convergence des plus grandes v.p. vers u_+? $\liminf_N \lambda_{max} \geq u_+$ p.s.
  Colle ou pas au reste du spectre? Espacement entre valeurs propres consécutives au bord.
  Même question pour les plus petites valeurs propres.

- Distribution limite des plus grandes v.p.?

- Universalité ou non de ces distributions limite? Classe d'universalité?

Le cas blanc: $\quad \Sigma = Id$.

L'ensemble de Wishart

Ensemble de Laguerre Unitaire (LUE) ou Orthogonal (LOE) =Wishart complexe ou réel.

$$\mu = \mathcal{N}(0,1).$$

Densité jointe des valeurs propres (Bronk (1965), Wishart (1928)):

$$f_L(\lambda_1,\ldots,\lambda_N) = \frac{1}{Z_N} \prod_{1\le i<j\le N} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^N \lambda_i^\alpha \exp\{-N\frac{\beta}{2}\lambda_i\},$$

avec $\alpha = p - N$ si $\beta = 2$ (complexe) et $\alpha = (p - N - 1)/2$ si $\beta = 1$ (réel).

Outils de calcul pour la loi des plus grandes valeurs propres: processus "determinantaux" dans le cas complexe (pfafffiens dans le cas réel)

Loi asymptotique des plus grandes valeurs propres

$$\gamma_N = \frac{p}{N}, \quad \rho_{Np} = \sigma^2 \left(1 + \gamma_N^{1/2}\right)^2, \quad \sigma_{Np} = \frac{\sigma^2}{N^{2/3}} \left(1 + \gamma_N^{1/2}\right) \left(1 + \gamma_N^{-1/2}\right)^{1/3}.$$

Théorème: TW (94), Johansson (2000) Johnstone (2001) El Karoui (2003)

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{\lambda_1 - \rho_{Np}}{\sigma_{Np}} \le x_1, \dots, \frac{\lambda_K - \rho_{Np}}{\sigma_{Np}} \le x_K \right) = F^K_{2(1)}(x_1, x_2, \dots, x_K),$$

avec $F^K_{2(1)}$ la loi de Tracy-Widom complexe (réelle) des K plus grandes v.p.

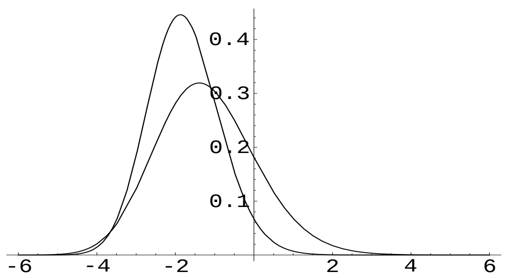


Figure 2: Courbe de gauche (resp. droite) : densité de la loi de TW complexe (resp. réelle) pour la plus grande valeur propre.

Conjecture d'universalité

Soit K un entier donné et $\Sigma = Id$. Si

$$\int x d\mu(x) = 0, \int |x|^2 d\mu(x) = \sigma^2 \text{ and } \int |x|^4 d\mu(x) < \infty, \tag{1}$$

alors on doit avoir: la loi asymptotique de $\left(\frac{\lambda_1^{sc} - \rho_{Np}}{\sigma_{Np}}, \frac{\lambda_2^{sc} - \rho_{Np}}{\sigma_{Np}}, \cdots, \frac{\lambda_K^{sc} - \rho_{Np}}{\sigma_{Np}}\right)$ est la loi de Tracy-Widom $F_{2(1)}^K$.

Théorème : Geman (80) Bai-Yin-Krishnaiah (88), Bai-Silverstein-Yin(88)
Si $\int |x|^4 d\mu(x) < \infty$, alors $\lim_{N\to\infty} \lambda_1 = u_+$ p.s.
Si $\int |x|^4 d\mu(x) = +\infty$, alors $\limsup_{N\to\infty} \lambda_1 = +\infty$ p.s.

Résultats d'universalité

Théorème à la Soshnikov S.P. (2007) $M_N = 1/NXX^*$. On suppose :

- μ est une loi symétrique $(\star)$.
- μ a des moments sous gaussiens (GD), i.e.

$$\exists C > 0,\ \forall k > 0,\ \int |x|^{2k} d\mu(x) \leq (Ck)^k.$$

- $\lim p/N \in [0, \infty]$.

Alors si K est un entier fixé

$$\left(\frac{\lambda_1^{sc} - \rho_{Np}}{\sigma_{Np}}, \frac{\lambda_2^{sc} - \rho_{Np}}{\sigma_{Np}}, \cdots, \frac{\lambda_K^{sc} - \rho_{Np}}{\sigma_{Np}} \right) \xrightarrow{d} F_{2(1)}^K.$$

La méthode des moments de Soshnikov

- LUE: $\lambda_1 = u_+ + C\xi N^{-2/3}$ avec $\xi \sim F_2^{TW}$

- Si on calcule
$$m_k^N(t_1,\ldots,t_k) = \mathbb{E}\prod_{i=1}^k Tr\left(\frac{M_N}{u_+}\right)^{[t_i N^{2/3}]},$$
pour tout k, en gros la transformée de Laplace de la loi jointe des plus grandes v.p.

- Pas de calcul exact de m_k^N,

1) montre qu'il est borné,

$$2) \quad |m_k^N(t_1,\ldots,t_k) - m_k^N(LUE)(t_1,\ldots,t_k)| = o(1).$$

L'approche de Tao et Vu

ThéorèmeTao-Vu (2009)
X_{ij} à decroissance sous exponentielle seulement. Alors universalité de la loi des K plus grandes valeurs propres.

Pas de condition de symétrie.

Idée: on regarde le comportement des statistiques locales de valeurs propres au bord du spectre:

$$\mathbb{E}f\left(N^{2/3}(\lambda_1 - u_+), \cdots, N^{2/3}(\lambda_K - u_+)\right).$$

On part d'une matrice de Wishart et on remplace une à une les entrées gaussiennes par des v.a. quelconques mais avec les mêmes premiers moments (4 en fait). Développement limité à la Lindebergh.

Autres résultats: Yau et al (2011). Approche par la résolvente.

Toujours pas la bonne condition sur les moments: 4, 36 ou 12 (Ruzmaikina (2006), Khorunzhy (2009))

Précisions à cette convergence

Théorème Ma (2008), Johnstone-Ma (2011) Pour l'ensemble de Wishart complexe (ou réel):

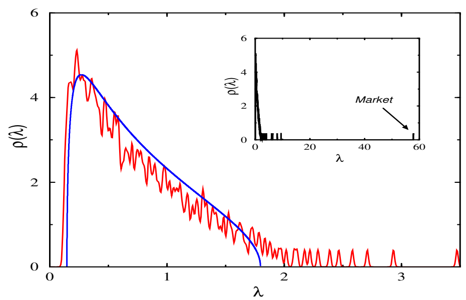
Pour tout s_0, $\exists N(s_o, \gamma)$ t.q. si $N, p \geq N(s_o, \gamma)$ alors $\forall s \geq s_0$,

$$\left| \mathbb{P}\left(\lambda_1 \leq \mu_{N,p} + \sigma_{N,p} s\right) - F_2^{TW}(s) \right| \leq C(s_0)(n \wedge p)^{-2/3} e^{-s/2},$$

avec C une fonction décroissante continue.

Basé sur la loi explicite de la plus grande v.p. sous forme d'un déterminant de Fredholm.
Qualité de l'approximation (même pour des matrices de taille "raisonnable").
Pour les lois autres que gaussiennes: Rien!

Le moment d'ordre 4.



Entrées à queue lourde

Soit $M = XX^*$ avec $p/N \to \gamma \in [1, \infty)$ tq. $\mathbb{E}(X_{ij}) = 0$ et

$$1 - F(x) = P(|X_{ij}| \geq x) = L(x)x^{-\alpha}, \quad 0 < \alpha < 4, \quad \forall i, j$$

où L est une fonction à variation lente, i.e., $\forall\, t > 0 \lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1$.
Soient $\lambda'_1 \geq \ldots \geq \lambda'_N$ les v.p. de M, $b_{Np} = \inf\{x : 1 - F(x) \leq \frac{1}{Np}\} \sim (Np)^{1/\alpha}$.

$$\mathcal{P}_N = \sum_{i=1}^{N} \delta_{b_{Np}^{-2}\lambda'_i}.$$

Théorème (Auffinger-Ben Arous-S. P.)(Soshnikov: Wigner, $\alpha < 2$).
$\mathcal{P}_N$ converge en loi, quand $N \to \infty$, vers un processus de Poisson $\mathcal{P}$ défini sur $(0, \infty)$ avec intensité $\rho(x) = \frac{\alpha}{2x^{1+\alpha/2}}$.

$$\Longrightarrow \lim_{N\to\infty} \mathbb{P}\left(\frac{1}{b_{Np}^2}\lambda'_1 \leq x\right) = \exp(-x^{-\frac{\alpha}{2}}).$$

Idée de la preuve.

$$\frac{1}{b_{Np}} \max_{1\le i\le N, 1\le j\le p} |X_{ij}| \text{ converge en loi.}$$

Les plus grandes v.p. sont déterminées par les plus grandes entrées de la matrice (A. Soshnikov (Wigner, $0 < \alpha < 2$).).

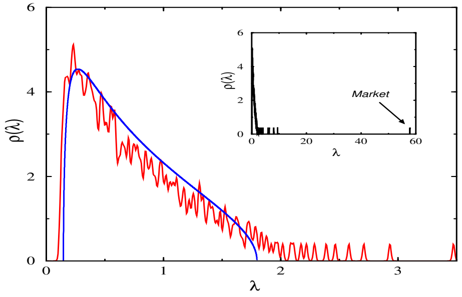
Si $0 < \alpha < 2$ les grands coeffs. sont suffisament grands.
Si $2 \le \alpha < 4$, "truc" de Biroli, Bouchaud and Potters (2006):

$$X = \left(X_{ij} 1_{|X_{ij}| \le N^\beta}\right) + \left(X_{ij} 1_{|X_{ij}| > N^\beta}\right),$$

pour un β bien choisi.

Borne le rayon spectral de $\left(X_{ij} 1_{|X_{ij}| \le N^\beta}\right)\left(X_{ij} 1_{|X_{ij}| \le N^\beta}\right)^*$ avec des outils standard de RMT et étude à part du rayon spectral de l'autre matrice (creuse).

Matrices non blanches.



Covariance plus complexe

"Spiked models": $\Sigma = \mathrm{diag}(\pi_1, \pi_2, \ldots, \pi_r, 1, \ldots, 1)$ perturbation de rang fini de l'identité, $\pi_1 > \pi_2 \geq \cdots \geq \pi_r > 1$.
Question: Impact des π_i sur les plus grandes v.p. de $M_N(\Sigma)$? λ_i estimateurs des π_i?

Théorème : Ensembles de Wishart Complexes. Baik-Ben Arous-P (2004), D. Paul (2005) $\sigma^2 = 1$

$$\text{Si } \pi_1 < 1 + \frac{1}{\sqrt{\gamma}}, \lambda_1 = u_+ + \frac{(1+\sqrt{\gamma})^{4/3}}{\gamma^{1/6} N^{2/3}} \xi^{TW}, \text{ fluctuations Tracy-Widom.}$$

$$\text{Si } \pi_1 > 1 + \frac{1}{\sqrt{\gamma}}, \lambda_1 = \gamma\pi_1 + \frac{\pi_1}{\pi_1 - 1} + \frac{1}{\sqrt{N}} G, \text{ où } G \text{ est une v.a. gaussienne}$$

$$\text{Si } \pi_1 = 1 + \frac{1}{\sqrt{\gamma}}, \text{ une autre loi limite.}$$

Remarques:

- Si $\pi_1 > 1 + \frac{1}{\sqrt{\gamma}}$ est de multiplicité $r > 1$, la loi limite de λ_1 est la loi de la plus grande v.p. d'un GUE $r \times r$ GUE (i.e. une matrice Hermitienne H avec H_{ij}, $i \leq j$ i.i.d gaussiennes standard).

- Σ diagonale sans importance ici (loi gaussienne).

- D. Paul idem pour des matrices réelles (plus dur!).
  Dans le cas complexe: calcul de la densité jointe des v.p. (James (1965) pour Σ quelconque).

$$f(\lambda_1, \ldots, \lambda_N) = 1/Z_N \det(e^{-N\lambda_j/\pi_i}) \prod_i \lambda_i^{p-N} V(\lambda), V : \text{ Vandermonde.}$$

Universalité : application en finances: secteurs (Bouchaud (2005)), génétique: structure (Patterson, Price, Reich (2006))

Extension à des ensembles non gaussiens

- Limite presque sûre de la plus grande v.p.:
  Théorème : Baik-Silverstein (2005).
  Hyp: $\int |x|^4 d\mu < \infty$. Alors,
  - Si $\pi_1 > 1 + \frac{1}{\sqrt{\gamma}}$ est de multiplicité $r \geq 1$, p.s.

$$\lim_{N\to\infty} \lambda_1 = \cdots = \lim_{N\to\infty} \lambda_r = \gamma\pi_1 + \frac{\pi_1}{\pi_1 - 1}.$$

  - Si $\pi_1 \leq 1 + \frac{1}{\sqrt{\gamma}}, \lim_{N\to\infty} \lambda_1 = u_+$.

- **Théorème :** Bai-Yao (2007), Féral-SP(2008).
  Hyp: $\int |x|^4 d\mu < \infty$ et $\pi_1 > 1 + \frac{1}{\sqrt{\gamma}}$. Le vecteur $\sqrt{N}(\lambda_i - \gamma\pi_1 - \frac{\pi_1}{\pi_1 - 1}), i \leq r$ converge en loi vers les v.p. d'une matrice aléatoire gaussienne, qui n'est en général PAS un GUE.

- Si $\pi_1 < 1 + \frac{1}{\sqrt{\gamma}}$ et μ vérifie ($\star$) et (GD) alors universalité.

Séparation de la plus grande v.p. dans le cas général

Théorème: Bai-Yao (2011)
Si $\pi_1 > z_c(H)$ avec $z_c(H)$ explicite on a que

$$\lambda_1 = \gamma\pi_1 + \int \frac{\pi_1}{\pi_1 - x} dH(x) + \frac{\chi}{\sqrt{N}},$$

avec χ une loi explicite (en général rien d'un GUE)

Autres déformations: W matrice déterministe et X à entrées $\mathcal{N}(0,1)$ complexes.

$$X \to X + W, \Sigma = Id,$$

Aussi une transition de phase: si $\lambda_{max}(W) > C(\gamma)$, $\lambda_1 > u_+$ avec des fluctuations dans une échelle $1/\sqrt{N}$.

Par analogie avec les matrices de Wigner (Capitaine-Donati-Féral (2008)), loi pourrait ne pas être universelle du tout.
Approche par les grandes déviations (Benaych Georges-Guionnet-Maida (2009, 2011))

Questions

Questions:

1. Comportement des plus grandes v.p. de $1/N(X-\overline{X})(X-\overline{X})^*$? ou encore en estimant la variance?
2. Vitesse de convergence vers la loi de Tracy-Widom ou autres distributions limites?
3. Moment d'ordre 4 est-il la bonne borne pour universalité ou pas?
4. Transition éventuelle entre Tracy-Widom et lois à queues lourdes?
5. Même transition de phase pour la plus petite valeur propre: universalité ou pas?

6. Application à de meilleurs estimateurs de Σ: les vecteurs propres de matrices aléatoires.
7. Quand rencontre-t-on Tracy-Widom?

Application à l'estimation de la covariance.

Quid des vecteurs propres

Comportement global : on sait peu de choses. Si X est une matrice Gaussienne: Si $\Sigma = Id$, soit (U, D) une diagonalisation de M_N: $M_N = UDU^*$ avec $U \in \mathbb{U}(N)$ et D une matrice diagonale réelle.

U est distribué suivant la mesure de Haar.

Preuve: Gram-Schmidt+invariance par rotation de la loi Gaussienne

Si $\Sigma \neq Id$, on ne sait RIEN.

Conjecture: si $\Sigma = Id$ et pour une matrice dont les entrées ne sont pas Gaussiennes avec $\mathbb{E}|X_{ij}|^4 < \infty$, la matrice des vecteurs propres de M_N doit être "asymptotiquement distribuée suivant la mesure de Haar".

Quelques résultats pour le comportement global: $\Sigma = Id$

Idée due à Silverstein ('95):
U est asymptotiquement distribué suivant la mesure de Haar si étant donné un vecteur $x \in \mathbb{S}^{N-1} = \{x \in \mathbb{R}^N, |x| = 1\}$ arbitraire, $y = Ux$ est asymptotiquement uniformément distribué sur la sphère unité.

Ou encore si on pose

$$Y_N(t) := \sqrt{\frac{N}{2}} \sum_{i=1}^{[Nt]} (|y_i|^2 - 1/N),$$

on a que $Y_N(t)$ converge en loi vers un Pont Brownien si y suit loi uniforme.

Théorème Bai-Miao-Pan (2007)
On suppose $\mathbb{E}|X_{ij}|^4 = 3$ (cas réel). Soit g une fonction analytique sur le support de la loi de Marchenko-Pastur ρ_{MP} Alors quand $N \to \infty$,

$$\int g(x) dG_N(x) \to \text{ une Gaussienne (centrée et de variance déterminée).}$$

Idée: $G_N(t) \simeq Y_N(F_N(t))$, $F_N(t)$: fonction de répartition empirique des v.p. de M_N.

Approche de Tao et Vu pour les matrices de Wigner

H_N dont les entrées sont à décroissance exponentielle et dont les 4 premiers moments sont les mêmes que le GOE (ou GUE).

Soient $(u_1, ..., u_N)$ ses vecteurs propres, de coordonnées $u_{i,j}, j = 1, \ldots, N$.
Soient $\xi_{i,p}, i \geq 1, p > 1$ i.i.d. $\mathcal{N}(0,1)$ et $\xi_{i,1} = |\mathcal{N}(0,1)|$.

Théorème Tao-Vu (2011)
$\exists \delta > 0$ tel que pour tout $k = O(N^\delta)$ et pour une fonction F bornée et "bien régulière"

$$\left| \mathbb{E}\left(F(\sqrt{N}(u_{i_a,j_b})_{1 \leq a,b \leq k} - F(\xi_{i_a,j_b})_{1 \leq a,b \leq k} \right) \right| = o(1).$$

Basé sur un résultat de Jiang (2006): une matrice orthogonale est presque une matrice gaussienne (si on considère un nombre $o(\sqrt{N})$ de ses colonnes).

Le cas $\Sigma \neq Id$

Fonctionelles du type:

$$\theta_N(g) := \frac{1}{N}\mathrm{Tr}\left(g(\Sigma)(M_N(\Sigma) - zI)^{-1}\right),$$

avec $z \in \mathbb{C}^+ = \{z \in \mathbb{C}, \Im z > 0\}$, $g(\Sigma) = V\mathrm{diag}(g(\pi_1), \ldots, g(\pi_N))V^*$ si V est la matrice des vecteurs propres de Σ avec g une fonction régulière (bornée avec nbr fini de discontinuités ou analytique).

Théorème: Ledoit-Péché (2009)
Si Supp$(H) \subset [h_1, h_2]$ avec $h_1 > 0$ et

$$\mathbb{E}|X_{ij}|^{12} < \infty \text{ indépendant de } N \text{ et } p.$$

Alors $\theta_N(g) \to \theta(g)$ p.s. quand $N \to \infty$ pour une fonction explicite (de H, γ).

Conséquences: on "connait" la moyenne de $N|u_i^* v_j|^2$ sur les vecteurs propres associés aux valeurs propres empiriques (resp. de la vraie covariance) dans l'intervalle $[\underline{\lambda}, \overline{\lambda}]$ (resp. $[\underline{\tau}, \overline{\tau}]$).

Un estimateur amélioré de la covariance

On cherche un estimateur de Σ sous la forme UD_NU^*, D_N diagonale i.e. un estimateur qui a les mêmes vecteurs propres que $M_N(\Sigma)$.
Le meilleur estimateur (norme de Frobenius) est

$$\widetilde{D}_N = \mathrm{diag}(\widetilde{d}_1, \ldots, \widetilde{d}_N) \quad \text{où} \quad \forall i = 1, \ldots, N \quad \widetilde{d}_i = u_i^* \Sigma_N u_i.$$

Il faut donc obtenir des informations sur les $\tilde{d}_i$.

On pose $\forall x \in \mathbb{R}$, $\quad \Delta_N(x) = \frac{1}{N} \sum_{i=1}^{N} \widetilde{d}_i \, 1_{[\lambda_i, +\infty)}(x) = \frac{1}{N} \sum_{i=1}^{N} u_i^* \Sigma_N u_i \times 1_{[\lambda_i, +\infty)}(x).$

Alors on a que $\forall i = 1, \ldots, N \qquad \widetilde{d}_i = \lim_{\varepsilon \to 0^+} \frac{\Delta_N(\lambda_i + \varepsilon) - \Delta_N(\lambda_i - \varepsilon)}{F_N(\lambda_i + \varepsilon) - F_N(\lambda_i - \varepsilon)}.$

$\Delta_N(x) \to \Delta(x)$. De plus $\Delta(x) = \int_{-\infty}^{x} \delta(\lambda) \, dF(\lambda)$

Ce qu'on gagne

On a effectué 10,000 simulations avec

$$\rho_N(\Sigma) = 0.2\delta_1 + 0.4\delta_3 + 0.4\delta_{10}, \qquad \gamma = 2$$

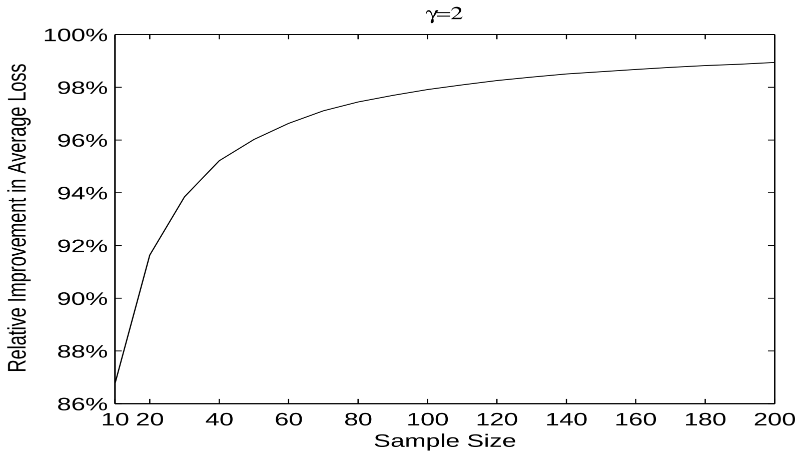
et en augmentant le nombre de variables p de 5 à 100. Pour chaque simulation, on a calculé le "Percentage Relative Improvement in Average Loss" (PRIAL):
si M est un estimateur de Σ_N et si $|A|_F^2 = \mathrm{Tr} AA^*$ (norme de Frobenius),

$$PRIAL(M) = 100 \times \left[1 - \frac{\mathbb{E} \left\| M - U_N \widetilde{D}_N U_N^* \right\|_F^2}{\mathbb{E} \left\| M_N(\Sigma) - U_N \widetilde{D}_N U_N^* \right\|_F^2} \right].$$

On a considéré $\widetilde{S}_N := UD'U^*$, où $D_i' = \lambda_i / |1 - \gamma^{-1} - \gamma^{-1} \lambda_i \breve{m}_\rho(\lambda_i)|^2$ qui est notre estimateur amélioré de Σ.

Simulations

On voit que même avec une petite taille, $p = 40$, on obtient un PRIAL de 95%.



Transition de phase pour les vecteurs propres

Théorème D. Paul (2006), X. Mestre (2009), Rao et Benaych-Georges (2010) Capitaine (2011).

Cas le plus simple:

$$\Sigma = \text{diag}(\pi_1, 1, \ldots, 1).$$

Soit u_1 (resp. e_1) vecteur propre normalisé de $M_N(\Sigma)$ (resp. de Σ) associé à λ_1 (resp. π_1):
Si $\pi_1 > 1 + 1/\sqrt{\gamma}$,

$$\lim_{N\to\infty} |< u_1, e_1 >| = \sqrt{\frac{1 - \gamma/(\pi_1 - 1)^2}{1 + \gamma/(\pi_1 - 1)}} \text{ a.s. } .$$

Si $\pi_1 \leq 1 + 1/\sqrt{\gamma}$,

$$\lim_{N\to\infty} |< u_1, e_1 >| = 0.$$

Cas général: limite p.s. de la projection de u_1 sur Ker$(\pi_1 - \Sigma)$ avec transition de phase.

Fluctuations plus précises des vecteurs propres?

Questions

- Estimateur amélioré de $\Sigma \neq Id$: convergence quand on remplace les $\tilde{d}_i$ par des estimateurs empiriques, un meilleur choix que garder les vecteurs propres de M_N?

- Comment se comportent le "reste" des vecteurs propres localisés dans le cas où il y a séparation des v.p.?

- Comment mélanger tout cela pour améliorer l'estimation de la covariance?

La loi de Tracy Widom et sa classe d'universalité.

La loi de Tracy-Widom complexe

Soit Ai la fonctiond'Airy standard et q la solution de l'équation de Painlevé II $\frac{\partial^2}{\partial x^2}q = xq(x) + 2q^3(x)$, avec $q(x) \sim Ai(x)$ quand $x \to +\infty$.
La loi de Tracy-Widom complexe (resp. réelle) de la plus grande v.p. est :

$$F_2^{TW}(x) = \exp\left\{\int_x^\infty (x-t)q^2(t)dt\right\} \quad \text{(GUE)}$$

$$F_1^{TW}(x) = \exp\left\{\int_x^\infty \frac{-q(t)}{2} + \frac{(x-t)}{2}q^2(t)dt\right\} \quad \text{(GOE).}$$

Dans le cas complexe on a aussi:

$$F_2^{TW}(x) = \det(I - K_{Ai})_{L^2([x,\infty[)} = \sum_k \frac{(-1)^k}{k!}\int_{[x,\infty)} \cdots \int_{[x,\infty)} \det K_{Ai}(x_i, x_j)\prod_{i=1}^k dx_i,$$

avec K_{Ai} le noyau d'Airy $K_{Ai}(x,y) = \frac{Ai(x)Ai'(y)-Ai'(x)Ai(y)}{x-y} = \int_0^\infty Ai(x+t)Ai(y+t)dt$.

D'où vient TW?

Loi de la plus grande v.p:

$$\mathbb{P}(\lambda_1 \leq x) = 1 - \mathbb{P}(\exists \text{ une vp dans } [x, \infty)) + \mathbb{P}(\exists \text{ 2 vp dans } [x, \infty)) \ldots$$

Inclusion Exclusion.

Calcul des probabilités de trouver k valeurs propres (non ordonnées) dans un intervalle donné: se fait via les fonctions de corrélation.

Définiton: Soit A un borélien de $\mathbb{R}^n$. On définit une mesure M_n par

$$M_n(A) = \mathbb{E}\left(\sum_{i_1 \neq i_2 \cdots \neq i_n} \mathbf{1}_A((x_{i_1}, \ldots, x_{i_n})) \right)$$

où les x_i sont les valeurs propres non ordonnées d'une matrice aléatoire. La densité de cette mesure / Lebesgue est la fonction de corrélation à n points $R_N^{(n)}$.

$R_N^{(n)}$: densité de valeurs propres (pas une densité de probabilité!)

Fonctions de corrélations pour LUE

On a aussi que si $f(x_1, \ldots, x_N)$ est la densité SYMÉTRISÉE des v.p

$$R_N^{(n)} = \frac{N!}{(N-n)!} \int_{\mathbb{R}^{N-n}} f(x_1, \ldots, x_N) \prod_{i=n+1}^{N} dx_i.$$

Pour le LUE (Wishart complexe), ce calcul est particulièrement simple:

$$f_L(\lambda_1, \ldots, \lambda_N) = \frac{1}{Z_N} (\det P_i(\lambda_j)_{i,j=1}^N)^2,$$

avec $P_i(x) = Q_i(x) x^{\alpha/2} e^{-Nx/2}$ Q_i polynôme orthogonal de Laguerre généralisé de degré $i-1$.

$$\int_{\mathbb{R}^+} Q_i(x) Q_j(x) x^\alpha e^{-Nx} dx = \delta_{ij}.$$

LUE et processus déterminantal

En utilisant les relations d'orthogonalité des polynômes généralisés de Laguerre, on montre que

$$R_N^{(n)}(x_1,\ldots,x_n) = \det K_N(x_i,x_j)_{i,j=1}^n,$$

avec

$$K_N(x,y) = C_N \frac{P_N(x)P_{N-1}(y) - P_N(y)P_{N-1}(x)}{x-y}.$$

Processus déterminantal.

En particulier $\mathbb{P}(\lambda_1 \le x) = \det(I - K_N)_{L^2(x,+\infty)}$.

Asymptotique: $x = u_+ + \frac{\alpha}{N^{2/3}}, y = u_+ + \frac{\beta}{N^{2/3}}$, convergence uniforme du noyau K_N vers le noyau d'Airy.

Occurrences de la loi de TW I

Ensembles invariants Un ensemble invariant complexe (resp. réel): loi $\mu^{(N)}$ sur les matrices Hermitiennes (res. réelles symétriques) tq

$$\mu^{(N)}(UMU^*) = \mu^{(N)}(M), \forall U \in \mathbb{U}(N) \text{ (resp. } U \in \mathbb{O}(N)).$$

Par exemple: $d\mu_N(H) = \frac{1}{Z_N} \exp\{-NV(H)\} \prod_{i<j} d\Re H_{ij} d\Im H_{ij} \prod_{i=1}^N dH_{ii}$, où V est un polynôme de degré pair.

Théorème: Deift et al. (1999, 2007)
$\exists\ C, u_+$ t.q.

$$CN^{2/3}(\lambda_1 - u_+) \to F_2^{TW}.$$

Structure de processus déterminantal (pol. orthogonaux).
Montre aussi que pour tous ces ensembles la mesure spectrale asymptotique a une densité en $\sqrt{u_+ - x}$.

Occurrences de la loi de TW II

Longueur de la plus longue sous suite croissante

$\mathcal{S}_N$ permutations de $\{1, 2, \ldots, \}$. Si $\Pi \in \mathcal{S}_N$ $\pi(i_1), \ldots, \pi(i_k)$ est une sous suite croissante si

$$i_1 < i_2 < \cdots < i_k \text{ et } \pi(i_1) < \pi(i_2) < \cdots < \pi(i_k).$$

$\mathcal{S}_N$ muni de la loi uniforme. l_N longueur plus longue sous suite croissante.

Théorème Baik-Deift-Johansson (1999)

$$\frac{l_N - 2\sqrt{N}}{N^{1/6}} \to F_2^{TW}.$$

Occurrences de la loi de TW III

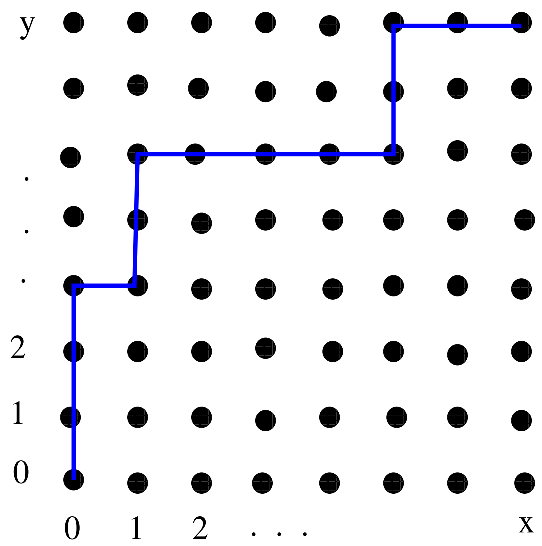
Modèles de percolation orientée

En chaque site $(i, j) \in \mathbb{N}^2$, on a une v.a. w_{ij} (temps d'attente). Les w_{ij} sont indépendants non nécessairement identiquement distribués.

Un chemin croissant π de $(0, 0)$ à $(x, y) \in \mathbb{N}^2$: une suite $(\pi_k \in \mathbb{Z}^2, k = 0, \ldots, x + y)$, avec $\pi_0 = (0, 0)$, $\pi_{x+y} = (x, y)$, et $\pi_{k+1} - \pi_k \in \{(1, 0), (0, 1)\}$.

Soit $L(\pi) = \sum_{(i,j)\in\pi} w_{i,j}$. Alors, le temps de dernier passage est défini par

$$G(x, y) = \max_{\pi:(0,0)\to(x,y)} L(\pi).$$



Modèles de percolation orientée II

Johansson (2000): Si la loi des $W_{i,j}$ est la loi exponentielle de paramètre 1, alors $G(x,y)$ a même loi que λ_{max} d'une matrice de Wishart complexe (non normalisée).

Conjecture: si $\mathbb{E}|W_{i,j}|^4 < \infty$ alors les fluctuations de $G(x,y)$ quand $x, y \to \infty$ avec $y/x \to \gamma$ sont TW.

$$m = \mathbb{E}W_{i,j}, \sigma^2 = \mathsf{Var}(W_{i,j}), \qquad \frac{G(x,y) - (x+y)m - 2\sigma\sqrt{xy}}{\sigma y^{-1/6} x^{1/2}} \to F_2^{TW}?$$

Théorème : Bodineau-Martin (2005) Baik-Suidan (2005)
Universalité si $\gamma = 0$ et $p = o(N^{\frac{3}{7}})$ et $\mathbb{E}|W_{ij}|^k < \infty \ \forall k$.

(Echelle en p plus restrictive si "moins de moments")

Pour une loi non exponentielle, aucune connexion connue à ce jour avec les matrices aléatoires.

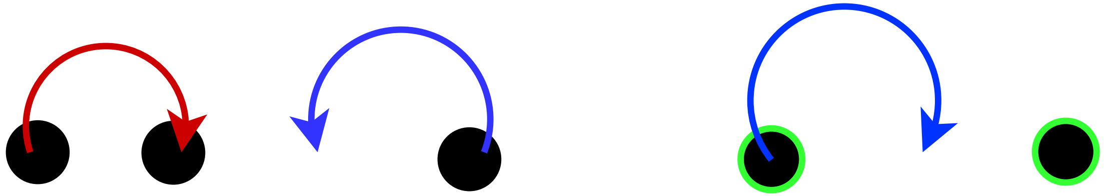
Occurrences de la loi de TW IV

Simple exclusion process Une configuration de particules $\eta_t \in \{0,1\}^{\mathbb{Z}}$, $t \geq 0$

$$\eta_t(i) = 1 : \text{ il y a une particule au site } i \text{ au temps } t;$$

Il y a au plus une particule en chaque site.

Étant donnée η_o, la dynamique est :
Les particules peuvent sauter seulement sur le site voisin (Simple) ssi ce site est vide (Exclusion). Une particule essaie de sauter avec proba p à droite (resp. $q = 1 - p$ à gauche).



Les sauts sont indépendants les uns des autres et arrivent après un temps d'attente exponentiel de taux 1.

Modèle de croissance associé au SEP

$$h_t(j) = \begin{cases} 2N_t + \sum_{i=1}^{j}(1 - 2\eta_i(t)), & \text{for } j \geq 1, \\ 2N_t, & \text{for } j = 0, \\ 2N_t - \sum_{i=j+1}^{0}(1 - 2\eta_i(t)), & \text{for } j \leq -1, \end{cases}.$$

où N_t est le nombre de particules qui ont sauté du site 0 au site 1 pendant l'intervalle de temps $[0, t]$.

Surface aléatoire "croissante": si $p = 1$ corner growth model.

Si $p = 1$ et $\eta_0 = 1_{\mathbb{Z}^-}$ alors connexion avec la plus grande valeur propre de matrices de Wishart complexes (Praehofer-Spohn (2001)). Connexion encore vraie si η_0: occupation des sites suit une loi de Bernoulli.

TW et fluctuations pour le SEP

Condition initiale de type pas: $\eta_0(i) = 1$ ssi $i \leq 0$.
Posons pour $v \in (0,1)$

$$H(s) = \frac{1+v^2}{2}t - s\frac{(1-v^2)^{2/3}}{2^{1/3}}t^{1/3}.$$

Théorème Johansson ('98) Tracy-Widom (2009) Soit $v \in [0,1)$ et $\gamma = p - q > 0$. On a

$$\lim_{t\to\infty} \mathbb{P}\left(h_{t/\gamma}(vt) \geq H(s)\right) = F_2^{TW}(s).$$

Si $0 < p < 1$ pas de processus déterminantal (série de formules magiques).

Différentes conditions initiales: Bernoulli (Borodin, Ferrari, S.P, Sasamoto, Spohn).

Loi de TW et répulsion

Répulsion des valeurs propres de matrices aléatoires.
Tasep, Asep: répulsion du fait de la règle d'exclusion.

Question: quelle intensité de répulsion?

Tasep i.e. $p = 1$: condition initiale de type "pas" ou "plate": loi de Tracy Widom complexe ou réelle.
Faut-il une répulsion de type "processus déterminantal"?
comment construire un processus déterminantal?
comment reconnaître un processus déterminantal?

Processus conditionnés à ne pas se croiser

Théorème Rains (2000)

Si une densité est de la forme

$$f(x_1, \ldots, x_N) = \frac{1}{Z} \det f_i(x_j) \det g_i(x_j),$$

alors elle induit un processus déterminantal (avec expression du noyau de corrélation).

Si $f_i(x) = p_t(y_i, x)$ est la probabilité de transition d'un processus Markovien, alors un théorème de Karlin-McGregor (1959) dit que

$$\det p_t(y_i, x_j) = \mathbb{P}(N \text{ processus aillent de Y à X sans se croiser}),$$

avec $Y = (y_1 < y_2 < \cdots < y_N)$ et $X = (x_1 < x_2 < \cdots < x_N)$.

Par exemple, LUE: carrés de Bessel conditionnés à ne pas se croiser; GUE: mouvements browniens.

Asep: un processus déterminantal ou pas?

T-W (2008) calcul de la probabilité de transition $P_Y(X,t)$ pour l'Asep. i.e la probabilité que N particules issues de $Y = (Y_1 < Y_2 < \cdots < Y_N)$ soient en $X = (X_1 < X_2 < \cdots < X_N)$ au temps t.

Théorème: Tracy-Widom (2008)

$$\mathbb{P}_Y(X,t) = \sum_{\sigma \in S_N} \oint \cdots \oint A_\sigma(\xi) \prod_{i=1}^{N} \xi_i^{x_i - y_{\sigma(i)} - 1} e^{tf(\xi_i)},$$

avec $f(\xi_i) = p/\xi_i + q\xi_i - 1$, et $A_\sigma = \prod_{\substack{i<j \\ \sigma(i)>\sigma(j)}} \left(-\frac{p+q\xi_{\sigma(i)}\xi_{\sigma(j)}-\xi_{\sigma(i)}}{p+q\xi_{\sigma(i)}\xi_{\sigma(j)}-\xi_{\sigma(j)}} \right)$.

Interpretation? Le terme : les processus de Poisson finissent bien où l'on veut. Le terme prend en compte le fait que des pas ont été supprimés pour que les trajectoires ne se croisent pas.

Après toute une série de formules magiques

Pour Asep avec condition initiale de type "pas", la loi de la particule la plus à droite est donnée par un déterminant de Fredholm...

Theorem: Tracy-Widom (2008)

$$\mathbb{P}(x_0(t) < x) = \det(I - K) := \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \oint \cdots \oint \det K(\xi_i, \xi_j)_{i,j=1}^{k},$$

pour un certain noyau K.

Cette formule est obtenue par une suite incroyable d'identités (déterminants de Cauchy....)
Pas d'interprétation en terme de processus déterminantal.

Pas de formule aussi simple pour la $m$tième particule pour $m \geq 1$.
Pas non plus de formule pour des conditions initiales autres que du type "pas".