Extreme values: a renormalization group approach

Eric Bertin

Laboratoire de Physique, ENS Lyon

Collaboration with G. Györgyi (Budapest), F. Angeletti and P. Abry (ENS Lyon)

Conference "Computation of transition trajectories and rare events in non-equilibrium systems", ENS Lyon, 11-15 June 2012

Many extreme value problems...

- Reaction path with the lowest barrier in a complex landscape
- Ground state in a disordered system
- Problems of pinned interfaces,...

In many cases, one needs to find minimum or maximum values among a set of random variables $\Rightarrow$ *statistics?*

See, e.g., Bouchaud, Mézard, J. Phys. A (1997).

Difficulties

- Presence of strong correlations, multiples scales,...
- Use of renormalization group could be relevant (but still difficult)
- What about the simplest extreme value problem, with iid random variables?

Standard results

- Variables $x_1, \ldots, x_n$ drawn from cumulative distribution $\mu(x)$ (called parent distribution)
- Rescaled cumulative distribution of $\max(x_1, \ldots, x_n)$

$$\mathcal{F}_\gamma(y) = exp[-(1+\gamma y)]^{-1/\gamma} \qquad 1+\gamma y > 0$$

$\gamma > 0$: Fréchet distribution (power-law tail of parent dist.)

$\gamma = 0$: Gumbel distribution (faster than powel-law tail)

$\gamma < 0$: Weibull distribution (bounded variables)

Fisher, Tippett (1928); Gnedenko (1943); Gumbel (1958)

Motivation

- Asymptotic distributions of extreme values of iid random variables known for long, but strong finite-size effects, not always easy to handle with standard probabilistic methods
- Idea: Use the renormalization language as a convenient tool to analyze fixed points and finite-size corrections, in spite of the absence of correlations
- Approach initiated in

  Györgyi, Moloney, Ozogány, Rácz, PRL (2008)

  Györgyi, Moloney, Ozogány, Rácz, Droz, PRE (2010)
- Aim of the present contribution: reformulate the results using a differential representation, which is more convenient

Basic renormalization idea

Extreme value statistics

- N iid random variables, cumulative distribution $\mu(x) = \int_{-\infty}^{x} \rho(x')dx'$
- Cumulative distribution for the maximum value

$$\mathrm{Prob}(\max(x_1, \ldots, x_N) < x) = \mathrm{Prob}(\forall i, x_i < x) = \mu^N(x)$$

Decimation procedure

- Split the set of sufficiently large N random variables x_i into $N' = N/p$ blocks of p random variables each
- y_j the maximum value in the j^{th} block

$$\max(x_1, \ldots, x_N) = \max(y_1, \ldots, y_{N'})$$

- y_j are also i.i.d. random variables, with a distribution $\mu_p(y)$

$$\mu_p(y) = \mu^p(y)$$

Renormalization operation

Raising to a power and rescaling

$$[\hat{R}_p \mu](x) = \mu^p\left(a_p x + b_p\right)$$

- Necessity of scale and shift parameters a_p and b_p to lift degeneracy of the distribution
- Conditions to fix a_p and b_p to be specified later on

Parameterization of the flow

- p considered as continuous rather than discrete
- change of flow parameter $p = e^s$: distribution $\mu(x,s)$, parameters $a(s)$ and $b(s)$
- Parent distribution $\mu(x)$ obtained for $s = 0$

$$\mu(x,0) = \mu(x)$$

Renormalization operation

Change of function

- double exponential form

$$\mu(x,s) = e^{-e^{-g(x,s)}}$$

- Link to the parent distribution: $g(x, s=0) = g(x)$

Standardization conditions

- Conditions to fix the parameters $a(s)$ and $b(s)$

$$\mu(0,s) \equiv e^{-1}, \qquad \partial_x \mu(0,s) \equiv e^{-1}$$

- In terms of the function $g(x,s)$

$$g(0,s) \equiv 0, \qquad \partial_x g(0,s) \equiv 1$$

Renormalization operation

Renormalization of $\mu(x, s)$

$$\mu(x, s) \equiv [\hat{R}_s \mu](x) = \mu^{e^s}\big(a(s)x + b(s)\big)$$

Renormalization of $g(x, s) = -\ln[-\ln \mu(x, s)]$

$$g(x, s) = g\big(a(s)x + b(s)\big) - s.$$

Very simple transformation: linear change of variable in the argument and global additive shift.

However, one needs to determine $a(s)$ and $b(s)$.

Differential representation

Iteration of the RG transformation

$$g(x, s + \Delta s) = [\hat{R}_{\Delta s} g](x, s)$$

Infinitesimal transformation $\Delta s = ds$

$$g(x, s + ds) = [\hat{R}_{ds} g](x, s)$$

- More explicitly, with $a(ds) = 1 + \gamma(s)ds$ and $b(ds) = \eta(s)ds$:

$$g(x, s + ds) = g\Big(\big(1 + \gamma(s)ds\big)x + \eta(s)ds, s\Big) - ds$$

  where the functions $\gamma(s)$ and $\eta(s)$ are to be specified

- Linearizing with respect to ds, we get

$$\partial_s g(x, s) = \big(\gamma(s)x + \eta(s)\big)\partial_x g(x, s) - 1$$

Partial differential equation

Determination of $\gamma(s)$ and $\eta(s)$

- Standardiz. conditions $g(0,s) \equiv 0$ and $\partial_x g(0,s) \equiv 1$ yield

$$\eta(s) \equiv 1$$

$$\gamma(s) = -\partial_x^2 g(0,s)$$

Partial differential equation of the flow

$$\partial_s g(x,s) = (1 + \gamma(s)\, x)\, \partial_x g(x,s) - 1$$

Fixed points of the flow

- Stationary solution $g(x,s) = f(x)$:

$$0 = (1+\gamma x) f'(x) - 1 \qquad \text{with} \ \ \gamma = -f''(0)$$

- Using the standardization condition $f(0) = 0$

$$f(x;\gamma) = \int_0^x (1+\gamma y)^{-1} dy = \frac{1}{\gamma} \ln(1+\gamma x)$$

- Fixed point integrated distribution

$$M(x;\gamma) = e^{-e^{-f(x;\gamma)}} = e^{-(1+\gamma x)^{-1/\gamma}}$$

Easy way to recover the well-known generalized extreme value distributions, obtained here as a fixed line of the RG transformation

Linear perturbations

- Perturbation $\phi(x,s)$ introduced through

$$g(x,s) = f(x) + f'(x)\,\phi(x,s)$$

- Linearized partial differential equation

$$\partial_s \phi(x,s) = (1+\gamma x)\,\partial_x \phi(x,s) - \gamma\,\phi(x,s) - x\,\partial_x^2 \phi(0,s)$$

- Convergence properties to the fixed point distribution are obtained from the analysis of this PDE

Eigenfunctions

Perturbations of the form $\phi(x,s) = e^{\gamma' s}\,\psi(x)$

Solution for the Weibull and Fréchet cases ($\gamma \neq 0$)

$$\psi(x;\gamma,\gamma') = \frac{1 + (\gamma'+\gamma)x - (1+\gamma x)^{\gamma'/\gamma+1}}{\gamma'(\gamma'+\gamma)}$$

in the range of x such that $1 + \gamma x > 0$.

Solution for the Gumbel case ($\gamma = 0$)

$$\psi(x;\gamma') = \frac{1}{\gamma'^2}\left(1 + \gamma' x - e^{\gamma' x}\right)$$

Empirical interpretation

- N variables in the block $\Rightarrow s = \ln N$
- Convergence $g(x, s = \ln N) \to f(x)$
- Corrections proportional to $e^{\gamma' s} \propto N^{\gamma'}$
  (if $\gamma' = 0$: logarithmic convergence in N).
- Interpretation of $\gamma' > 0$? Are there unstable solutions?

  $\Rightarrow$ Can we look at non-perturbative solutions?

Motivation

Unstable solutions around the fixed point may seem counterintuitive: can we find an example of full RG trajectory starting from an unstable direction?

Back to the equations: the Gumbel case

- Equation to be solved

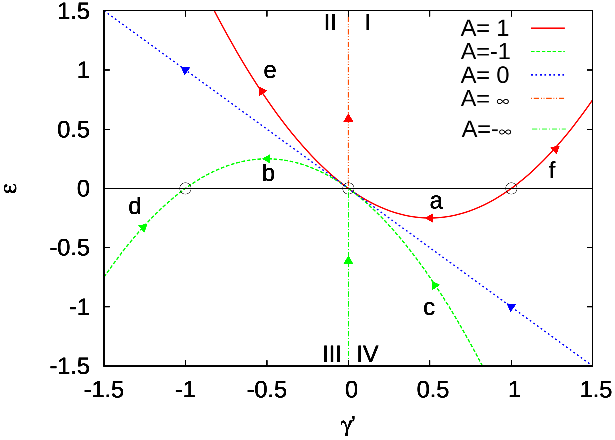
$$\partial_s g(x,s) = (1+\gamma(s)\,x)\,\partial_x g(x,s) - 1$$

- Ansatz for the solution starting from $f(x) = x$

$$g(x,s) = x + \epsilon(s)\,\psi\big(x;\gamma'(s)\big)$$

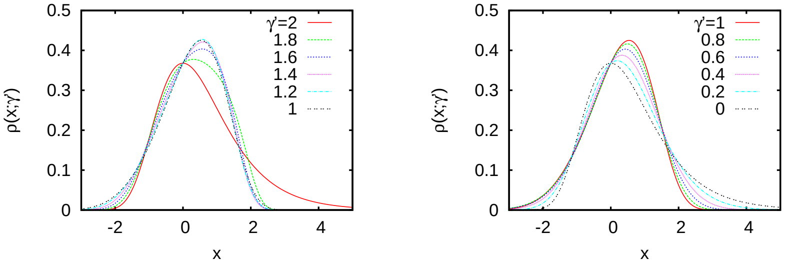
- Same as linear perturbation, except that γ' depends on s

Parameter space (ϵ, γ')



Evolution of the distributions

Starting close to the Gumbel distribution ($\gamma' = 2$)… and coming back to it (at $\gamma' = 0$) after an excursion



Bertin, Györgyi, J. Stat. Mech. (2010)

Generalization of standard extreme statistics

Raising the variables to an increasing power

- Choose iid variables whose statistics depends on the sample size n, for instance by raising $x_1, \ldots, x_n$ to a power q_n
- Question: statistics of the quantity $\max(x_1^{q_n}, \ldots, x_n^{q_n})$
- Motivation: link with the Random Energy Model

  Ben Arous et.al. (2005), Bogachev (2007)

Results

- Emergence of new limit distributions, for $q(n) \sim n^Q$

$$\mathcal{F}_{\gamma,Q} = \exp\left[-\left(1 - \frac{Q}{\gamma} \ln(1+\gamma x)\right)^{1/Q} \right]$$

- Standard distributions recovered for $Q \to 0$

Angeletti, Bertin, Abry, J. Phys. A (2012)

Formal analogy between sums and extremes

Extreme value statistics for iid random variables

- Relevant mathematical object: integrated distribution $\mu(x)$
- Integrated distribution of the maximum of N iid random variables

$$\mu_N(x) = \mu(x)^N$$

- Linear rescaling of x to preserve the standardiz. conditions

Statistics of sums of iid random variables

- Relevant mathematical object: characteristic function $\Phi(q)$
- Characteristic function for the sum of N iid random variables

$$\Phi_N(q) = \Phi(q)^N$$

- Linear rescaling

Same formal structure, only the objects differ

Result for the characteristic function

$$\Phi(q;\gamma) = e^{-|q|^{-\frac{1}{\gamma}}}$$

- Characteristic function of the symmetric Lévy distribution, of parameter $\alpha = -1/\gamma$.
- Here, one restriction: $\gamma \leq -\frac{1}{2}$, equivalent to $0 < \alpha \leq 2$
- Linear stability analysis (eigenfunctions, ...) can be performed in the same way as for extreme value statistics

Bertin, Györgyi, J. Stat. Mech. (2010)

Conclusion

On the present work

- Renormalization is a convenient tool to analyze fixed points and finite size corrections
- Analysis of finite size corrections made easy by the use of eigenfunctions
- Can be applied to variants of the present problems, for instance, statistics of $\max(x_1^{q_n}, \ldots, x_n^{q_n})$

Outlook

- Is renormalization without correlation really renormalization? Extension to correlated variables welcome... but yet unclear