Weighted Covariance and Correlation Matrices: Old Friends and Power-Estimators

David Morton de Lachapelle[1,2], Olivier Lévèque[2]

[1]Swissquote Bank (www.swissquote.ch), [2]EPFL-Lausanne (CH)

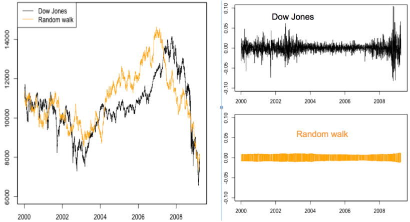
Workshop on Random Matrices, Information Theory and Applications
(EPFL-UMLV, Paris, 2009)

Outline

1. Weighting market events in time, why, how?
2. Sample Weighted Covariance Matrices
3. Asymptotic Weighted Correlation Matrices: spectral properties
4. Some results on real data
5. Conclusion

Motivations

- Time weightings improves the quality of volatility-based forecasts (covariance, variance, Value-at-Risk (VaR), ...).
- E.g.: Use decreasing weights to take advantage of *volatility clustering*.

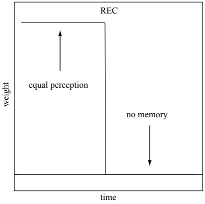


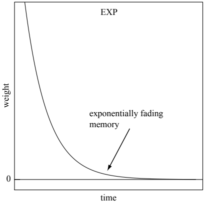
- Time weightings embody the limited and fading memory of market participants.

SWISSQUOTE

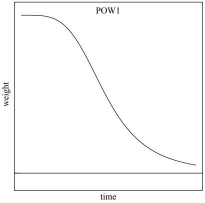
Weight profiles: the shape of memory

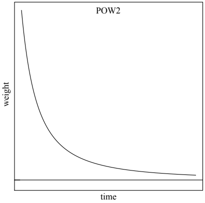
- Extreme cases: uniform (**REC**) and exponential (**EXP**) weightings





- Do markets forget all about their past? Sparing long-term memory.





Weight profiles: definitions

REC	EXP
$w(k) = \frac{1}{T} 1_{\{1 \leq k \leq T\}}$	$w(k) \sim \left(1 - \frac{1}{c}\right)^k$

- Power-law decay of memory:

POW1	POW2
$w(k) \sim \frac{1}{1+\left(\frac{k}{c}\right)^\gamma}$	$w(k) \sim \frac{1}{\left(1+\frac{k}{c}\right)^\gamma}$

- As $\gamma \to \infty$, POW1$\sim$REC and POW2$\sim$EXP.

Weight profiles: their attributes

- Attributes are useful for
  - understanding the role of parameters.
  - comparing profiles with each other.

		REC	EXP	POW1	POW2
$\bar{t}$	$\int_0^\infty t\,\alpha(t)\,dt$	$\frac{T}{2}$	c	$c\,(2\cos\frac{\pi}{\gamma})^{-1}$ $(\gamma > 2)$	$c\,(\gamma-2)^{-1}$ $(\gamma > 2)$
d	$\int_0^d \alpha(t)\,dt = 1-\varepsilon$	$\begin{cases}(1-\varepsilon)T & d \le T \\ 0 & d > T\end{cases}$	$c\,\log(\frac{1}{\varepsilon})$	$\sim c\,(\frac{1}{\varepsilon})^{\frac{1}{\gamma}-1}$	$c\left((\frac{1}{\varepsilon})^{\frac{1}{\gamma-1}}-1\right)$
$\delta_{0.5}$	$\frac{\alpha(\delta_{0.5})}{\alpha(0)} = 0.5$	/	$c\,\log 2$	c	$c\left(2^{\frac{1}{\gamma}}-1\right)$

SWISSQUOTE

Definition

The $N \times N$ sample weighted covariance matrix of returns:

$$\Sigma_{ij} = \frac{1}{N} \sum_{k=0}^{T} w_N(k) h_{ik} h_{jk},$$

where h_{ik} = return of asset i at time k, $w_N(k) \geq 0$ and $\frac{1}{N} \sum w_N(k) = 1$).
N = number of assets.

- Σ embeds volatility and correlation risk.
- Important forecaster in
  - risk assessment (e.g. volatility, value-at-risk),
  - optimization (e.g. portfolio allocation, trading algorithms),
  - product pricing (e.g. options, baskets).

EPFL
SWISSQUOTE

Definition

The $N \times N$ sample weighted covariance matrix of returns:

$$\Sigma_{ij} = \frac{1}{N} \sum_{k=0}^{T} w_N(k) h_{ik} h_{jk},$$

where $h_{ik} = \text{return of asset } i \text{ at time } k$, $w_N(k) \geq 0$ and $\frac{1}{N} \sum w_N(k) = 1$).
$N = \text{number of assets}$.

- Σ embeds volatility and correlation risk.
- Important forecaster in
  - risk assessment (e.g. volatility, value-at-risk),
  - optimization (e.g. portfolio allocation, trading algorithms),
  - product pricing (e.g. options, baskets).

EPFL SWISSQUOTE

Definition

The $N \times N$ sample weighted covariance matrix of returns:

$$\Sigma_{ij} = \frac{1}{N} \sum_{k=0}^{T} w_N(k) h_{ik} h_{jk},$$

where h_{ik} = return of asset i at time k, $w_N(k) \geq 0$ and $\frac{1}{N}\sum w_N(k) = 1$).
N = number of assets.

- Σ embeds volatility and correlation risk.
- Important forecaster in
  - risk assessment (e.g. volatility, value-at-risk),
  - optimization (e.g. portfolio allocation, trading algorithms),
  - product pricing (e.g. options, baskets).

Dynamics of Σ

- Σ_{ij} seen as the conditional covariance at $k+1$ $(k \in \{0, 1, \dots, T\})$:

$$\Sigma_{ij}(k+1) = \sum_{\ell=0}^{k} w_N(\ell) h_{i\ell} h_{j\ell},$$

$$\underset{\Sigma(0)=0}{\Longleftrightarrow}$$

$$\Sigma_{ij}(k+1) = w_N(0) h_{ik} h_{jk} + \frac{w_N(T-k+1)}{w_N(T-k)} \Sigma_{ij}(k)$$

- Any weighted covariance matrix can be uniquely decomposed as a contemporaneous contribution from the returns plus a term of conditional covariance.
- Close to the famous GARCH(1,1) in econometrics (Bollerslev, 1986).

EPFL SWISSQUOTE

Dynamics of Σ

- Σ_{ij} seen as the conditional covariance at $k+1$ ($k \in \{0, 1, \ldots, T\}$):

$$\Sigma_{ij}(k+1) = \sum_{\ell=0}^{k} w_N(\ell) h_{i\ell} h_{j\ell},$$

$$\underset{\Sigma(0)=0}{\Longleftrightarrow}$$

$$\Sigma_{ij}(k+1) = w_N(0) h_{ik} h_{jk} + \frac{w_N(T-k+1)}{w_N(T-k)} \Sigma_{ij}(k)$$

- Any weighted covariance matrix can be uniquely decomposed as a contemporaneous contribution from the returns plus a term of conditional covariance.
- Close to the famous GARCH(1,1) in econometrics (Bollerslev, 1986).

EPFL
SWISSQUOTE

Dynamics of Σ

- Σ_{ij} seen as the conditional covariance at $k+1$ ($k \in \{0, 1, \ldots, T\}$):

$$\Sigma_{ij}(k+1) = \sum_{\ell=0}^{k} w_N(\ell) h_{i\ell} h_{j\ell},$$

$$\underset{\Sigma(0)=0}{\Longleftrightarrow}$$

$$\Sigma_{ij}(k+1) = w_N(0) h_{ik} h_{jk} + \frac{w_N(T-k+1)}{w_N(T-k)} \Sigma_{ij}(k)$$

- Any weighted covariance matrix can be uniquely decomposed as a contemporaneous contribution from the returns plus a term of conditional covariance.
- Close to the famous GARCH(1,1) in econometrics (Bollerslev, 1986).

Weighted volatility processes

- We set $N = 1$ and define the following stochastic process

$$h_k = \sigma_k \varepsilon_k, \text{ where } \varepsilon_k \sim \text{i.i.d.}, E(\varepsilon_k) = 0, E(\varepsilon^2) = 1.$$

The conditional volatility obeys

$$\sigma^2(k+1) = w(0)h_k^2 + \frac{w(T-k+1)}{w(T-k)}\sigma^2(k), \; k \in \{0, 1, \ldots, T\}.$$

- No need for the distribution of returns, only their unconditional distribution.
- In general the process is non-stationary and $E(h_k^{2m})$ may not exist.

Weighted volatility processes

- We set $N = 1$ and define the following stochastic process

$$h_k = \sigma_k \varepsilon_k, \text{ where } \varepsilon_k \sim \text{i.i.d.}, E(\varepsilon_k) = 0, E(\varepsilon^2) = 1.$$

The conditional volatility obeys

$$\sigma^2(k+1) = w(0)h_k^2 + \frac{w(T-k+1)}{w(T-k)}\sigma^2(k), \ k \in \{0, 1, \dots, T\}.$$

- No need for the distribution of returns, only their unconditional distribution.
- In general the process is non-stationary and $E(h_k^{2m})$ may not exist.

Weighted volatility processes

- We set $N = 1$ and define the following stochastic process

$$h_k = \sigma_k \varepsilon_k, \text{ where } \varepsilon_k \sim \text{i.i.d.}, E(\varepsilon_k) = 0, E(\varepsilon^2) = 1.$$

The conditional volatility obeys

$$\sigma^2(k+1) = w(0)h_k^2 + \frac{w(T-k+1)}{w(T-k)}\sigma^2(k), \; k \in \{0, 1, \ldots, T\}.$$

- No need for the distribution of returns, only their unconditional distribution.
- In general the process is non-stationary and $E(h_k^{2m})$ may not exist.

Example: IGARCH(1), EWMA

- The linear IGARCH(1) (Engle and Bollerslev, 1986), or Exponentially Weighted Moving Average (EWMA) (RiskMetrics ,1996) follows from the choice $w(k) \sim (1+\frac{1}{c})^k$:

$$\sigma^2(k+1) = \frac{1}{c} h_k^2 + \left(1+\frac{1}{c}\right) \sigma^2(k).$$

- Interestingly, linear IGARCH(1) is strongly stationary but not weakly stationary (even moments are not defined in the limit $k \to \infty$) (Nelson, 1990).

EPFL
SWISSQUOTE

Example: IGARCH(1), EWMA

- The linear IGARCH(1) (Engle and Bollerslev, 1986), or Exponentially Weighted Moving Average (EWMA) (RiskMetrics ,1996) follows from the choice $w(k) \sim (1 + \frac{1}{c})^k$:

$$\sigma^2(k+1) = \frac{1}{c} h_k^2 + \left(1 + \frac{1}{c}\right) \sigma^2(k).$$

- Interestingly, linear IGARCH(1) is strongly stationary but not weakly stationary (even moments are not defined in the limit $k \to \infty$) (Nelson, 1990).

- The covariance decomposition is very useful for working out statistical properties of the process.
- Taking $\varepsilon_k \sim N(0,1)$ and $\sigma_0 = \text{cst.}$, and assuming the existence of the second and forth moment, the (non-stationary) kurtosis reads

$$\mathrm{Kurt}(h_k) = \frac{\mathrm{E}(h_k^4)}{\mathrm{E}(h_k^2)^2} = 3\prod_{i=0}^{k-1}\left(1 + \frac{2w(0)^2}{(w(0)+f(i))^2}\right) > 3,$$

where $f(i) = w(T-i+1)/w(T-i)$ is the ith weight increment.
- **Conclusion**: weighted-volatility processes generate excess kurtosis for all h_k.
- E.g.: $\mathrm{Kurt}(h_k^{EWMA}) = 3(1+\frac{2}{c^2})^{k-1}$, which diverges exponentially fast as $k, T \to \infty$.

Weighted-volatility excess kurtosis

- The covariance decomposition is very useful for working out statistical properties of the process.
- Taking $\varepsilon_k \sim N(0,1)$ and $\sigma_0 = \text{cst.}$, and assuming the existence of the second and forth moment, the (non-stationary) kurtosis reads

$$\text{Kurt}(h_k) = \frac{\mathrm{E}(h_k^4)}{\mathrm{E}(h_k^2)^2} = 3\prod_{i=0}^{k-1}\left(1 + \frac{2w(0)^2}{(w(0)+f(i))^2}\right) > 3,$$

where $f(i) = w(T-i+1)/w(T-i)$ is the ith weight increment.

- **Conclusion**: weighted-volatility processes generate excess kurtosis for all h_k.
- E.g.: $\text{Kurt}(h_k^{EWMA}) = 3(1+\frac{2}{c^2})^{k-1}$, which diverges exponentially fast as $k, T \to \infty$.

Weighted-volatility excess kurtosis

- The covariance decomposition is very useful for working out statistical properties of the process.
- Taking $\varepsilon_k \sim N(0,1)$ and $\sigma_0 = \text{cst.}$, and assuming the existence of the second and forth moment, the (non-stationary) kurtosis reads

$$\text{Kurt}(h_k) = \frac{\text{E}(h_k^4)}{\text{E}(h_k^2)^2} = 3\prod_{i=0}^{k-1}\left(1 + \frac{2w(0)^2}{(w(0)+f(i))^2}\right) > 3,$$

where $f(i) = w(T-i+1)/w(T-i)$ is the ith weight increment.

- **Conclusion**: weighted-volatility processes generate excess kurtosis for all h_k.
- E.g.: $\text{Kurt}(h_k^{EWMA}) = 3(1+\frac{2}{c^2})^{k-1}$, which diverges exponentially fast as $k, T \to \infty$.

Weighted-volatility excess kurtosis

- The covariance decomposition is very useful for working out statistical properties of the process.
- Taking $\varepsilon_k \sim N(0,1)$ and $\sigma_0 = \text{cst.}$, and assuming the existence of the second and forth moment, the (non-stationary) kurtosis reads

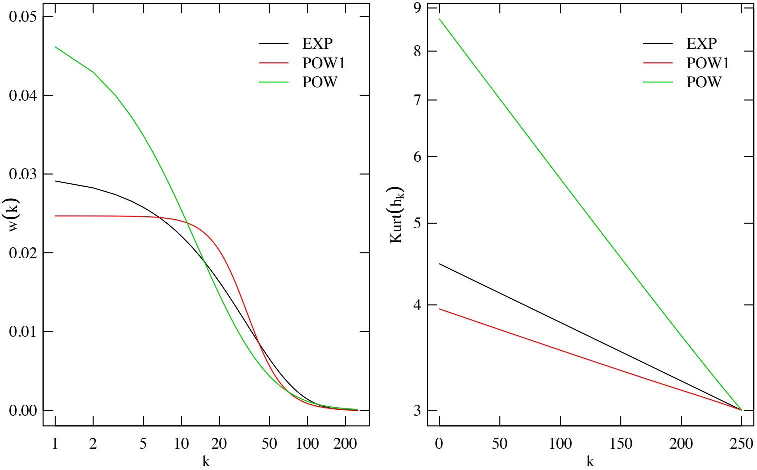
$$\text{Kurt}(h_k) = \frac{\text{E}(h_k^4)}{\text{E}(h_k^2)^2} = 3\prod_{i=0}^{k-1}\left(1 + \frac{2w(0)^2}{(w(0)+f(i))^2}\right) > 3,$$

where $f(i) = w(T-i+1)/w(T-i)$ is the ith weight increment.

- **Conclusion**: weighted-volatility processes generate excess kurtosis for all h_k.
- E.g.: $\text{Kurt}(h_k^{EWMA}) = 3(1+\frac{2}{c^2})^{k-1}$, which diverges exponentially fast as $k, T \to \infty$.

Weighted-volatility excess kurtosis: plots

- Comparison of excess kurtosis across profiles



EPFL
SWISSQUOTE

The sample autocorrelation function

- The two-point autocorrelation function of the squared returns can be calculated in closed-form. The general form is complicated, but for EWMA:

$$\rho(h_k^2, h_{k-\ell}^2) = \frac{\mathrm{E}(h_k^2, h_{k-\ell}^2) - \mathrm{E}(h_k^2)\mathrm{E}(h_{k-\ell}^2)}{\sqrt{\mathrm{V}h_k^2}\sqrt{\mathrm{V}h_{k-\ell}^2}} \sim (1+2\alpha^2)^{-\ell/2}, \; k \gg \ell,$$

  as previously found by Ding and Granger (1996).

- When $w(0)^2 \ll (w(0)+f(i))^2$ holds (consistent with GARCH(1,1)), we have

$$\rho(h_k^2, h_{k-\ell}^2) \sim \prod_{i=k-\ell}^{k-1} (w(0)+f(i))^2 \geq \prod_{i=k-\ell}^{k-1} f(i)^2 = \left(\frac{w(k+l+1)}{w(k+1)}\right)^2.$$

The sample autocorrelation function

- The two-point autocorrelation function of the squared returns can be calculated in closed-form. The general form is complicated, but for EWMA:

$$\rho(h_k^2, h_{k-\ell}^2) = \frac{\mathrm{E}(h_k^2, h_{k-\ell}^2) - \mathrm{E}(h_k^2)\mathrm{E}(h_{k-\ell}^2)}{\sqrt{\mathrm{V}h_k^2}\sqrt{\mathrm{V}h_{k-\ell}^2}} \sim (1+2\alpha^2)^{-\ell/2}, \; k \gg \ell,$$

  as previously found by Ding and Granger (1996).
- When $w(0)^2 \ll (w(0)+f(i))^2$ holds (consistent with GARCH(1,1)), we have

$$\rho(h_k^2, h_{k-\ell}^2) \sim \prod_{i=k-\ell}^{k-1} (w(0)+f(i))^2 \geq \prod_{i=k-\ell}^{k-1} f(i)^2 = \left(\frac{w(k+l+1)}{w(k+1)}\right)^2.$$

EPFL
SWISSQUOTE

Weighted sample correlation matrices

- The weighted sample correlation matrices is defined as

$$C = \frac{1}{N} H \mathrm{diag}(w_N) H^t,$$

  $H \in \mathbb{R}^{N \times T}$ is the matrix of centered (i.e. $\mu_i = 0$) and standardized (i.e. $\sigma_i = 1$) returns. Weights are normalized ($\frac{1}{N} \sum w_N = 1$).
- **Goal**: Find the **spectral density** $p(\lambda)$ and the **edge spectrum** $\{\lambda_{\mathrm{min}}, \lambda_{\mathrm{max}}\}$ of C for h_{ik} i.i.d. random variables with zero mean and unit variance (null model).
- Calculations are done in the two asymptotic limits $T/N = c_0 < \infty$ and $T/N = \infty$.

EPFL SWISSQUOTE

Weighted sample correlation matrices

- The weighted sample correlation matrices is defined as

$$C = \frac{1}{N} H \mathrm{diag}(w_N) H^t,$$

  $H \in \mathbb{R}^{N\times T}$ is the matrix of centered (i.e. $\mu_i = 0$) and standardized (i.e. $\sigma_i = 1$) returns. Weights are normalized ($\frac{1}{N}\sum w_N = 1$).
- **Goal**: Find the **spectral density** $p(\lambda)$ and the **edge spectrum** $\{\lambda_{\min}, \lambda_{\max}\}$ of C for h_{ik} i.i.d. random variables with zero mean and unit variance (null model).
- Calculations are done in the two asymptotic limits $T/N = c_0 < \infty$ and $T/N = \infty$.

Weighted sample correlation matrices

- The weighted sample correlation matrices is defined as

$$C = \frac{1}{N} H\text{diag}(w_N)H^t,$$

$H \in \mathbb{R}^{N \times T}$ is the matrix of centered (i.e. $\mu_i = 0$) and standardized (i.e. $\sigma_i = 1$) returns. Weights are normalized ($\frac{1}{N}\sum w_N = 1$).

- **Goal**: Find the **spectral density** $p(\lambda)$ and the **edge spectrum** $\{\lambda_{\text{min}}, \lambda_{\text{max}}\}$ of C for h_{ik} i.i.d. random variables with zero mean and unit variance (null model).
- Calculations are done in the two asymptotic limits $T/N = c_0 < \infty$ and $T/N = \infty$.

EPFL
SWISSQUOTE

- Rigorous, powerful (non-normal, non-i.i.d. returns).
- Brought to finance in 1999 by Laloux, Cizeau, Bouchaud, and Potters (other approaches: R-transform, Replica, ...).
- Derives an equation for the Stieltjes transform $g(z)$ of $p(\lambda)$ when $T/N \to c_0 < \infty$.
- The result extends to weighted estimators:

$$g(z) = \left(\int_0^{c_0} \frac{\alpha(t)}{1+\alpha(t)g(z)} dt - z \right)^{-1},$$

where $\alpha(t) = \lim_{N\to\infty} w_N(\lfloor Nt \rfloor)$, $\forall t \in [0, c_0]$.

Marčenko and Pastur approach

- Rigorous, powerful (non-normal, non-i.i.d. returns).
- Brought to finance in 1999 by Laloux, Cizeau, Bouchaud, and Potters (other approaches: R-transform, Replica, ...).
- Derives an equation for the Stieltjes transform $g(z)$ of $p(\lambda)$ when $T/N \to c_0 < \infty$.
- The result extends to weighted estimators:

$$g(z) = \left(\int_0^{c_0} \frac{\alpha(t)}{1+\alpha(t)g(z)} dt - z \right)^{-1},$$

where $\alpha(t) = \lim_{N\to\infty} w_N(\lfloor Nt \rfloor)$, $\forall t \in [0, c_0]$.

Marčenko and Pastur approach

- Rigorous, powerful (non-normal, non-i.i.d. returns).
- Brought to finance in 1999 by Laloux, Cizeau, Bouchaud, and Potters (other approaches: R-transform, Replica, ...).
- Derives an equation for the Stieltjes transform $g(z)$ of $p(\lambda)$ when $T/N \to c_0 < \infty$.
- The result extends to weighted estimators:

$$g(z) = \left(\int_0^{c_0} \frac{\alpha(t)}{1+\alpha(t)g(z)} dt - z \right)^{-1},$$

where $\alpha(t) = \lim_{N\to\infty} w_N(\lfloor Nt \rfloor)$, $\forall t \in [0, c_0]$.

Marčenko and Pastur approach

- Rigorous, powerful (non-normal, non-i.i.d. returns).
- Brought to finance in 1999 by Laloux, Cizeau, Bouchaud, and Potters (other approaches: R-transform, Replica, ...).
- Derives an equation for the Stieltjes transform $g(z)$ of $p(\lambda)$ when $T/N \to c_0 < \infty$.
- The result extends to weighted estimators:

$$g(z) = \left(\int_0^{c_0} \frac{\alpha(t)}{1 + \alpha(t) g(z)} dt - z \right)^{-1},$$

where $\alpha(t) = \lim_{N\to\infty} w_N(\lfloor Nt \rfloor)$, $\forall t \in [0, c_0]$.

The limit $T/N \to \infty$

- The limit $T/N \to \infty$ often leads to simpler calculations. Does MP extend to this limit?

Theorem

If $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous and decreasing function such that $\alpha \in L_2(\mathbb{R}^+)$, then

$$G_z(g) = g - \left(\int_0^\infty \frac{\alpha(t)}{1+\alpha(t)g} dt - z \right)^{-1}$$

admits a unique zero g^* that is the Stieltjes transform of a distribution.

Proof.

Show that $G_z(g)$ is a contraction on $\mathbb{C}_{++} = \{g \in \mathbb{C} : \text{Re } g \geq 0, \text{Im } g \geq 0\}$ (M. de Lachapelle, Lévèque, 2009).

The limit $T/N \to \infty$

- The limit $T/N \to \infty$ often leads to simpler calculations. Does MP extend to this limit?

Theorem

If $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ *is a continuous and decreasing function such that* $\alpha \in L_2(\mathbb{R}^+)$*, then*

$$G_z(g) = g - \left(\int_0^\infty \frac{\alpha(t)}{1+\alpha(t)g} dt - z \right)^{-1}$$

admits a unique zero g^* *that is the Stieltjes transform of a distribution.*

Proof.

Show that $G_z(g)$ is a contraction on $\mathbb{C}_{++} = \{g \in \mathbb{C} : \text{Re } g \geq 0, \text{Im } g \geq 0\}$ (M. de Lachapelle, Lévèque, 2009). □

- $G_z(g)$ admits a unique zero and is holomorphic. The Newton-Raphson method is simple and efficient, but requires

$$\lim_{\varepsilon \to 0^+} G_{\lambda+i\varepsilon}(g) = G_\lambda(g),$$

which is garanteed by Silverstein and Choi, 1995.

- **Require:** λ, $k_{\max}$, tol

  $g_0 \leftarrow$ random starter in $\mathbb{C} \setminus \mathbb{R}$
  for $k = 1$ to $k_{\max}$ **do**
  $g_k \leftarrow g_{k-1} - G_\lambda(g_{k-1})/G'_\lambda(g_{k-1})$
  if $|g_k - g_{k-1}| \leq$ tol **then**
  $g^* \leftarrow g_k$
  exit loop
  end if
  end for
  $p(\lambda) \leftarrow \frac{1}{\pi}\text{Im } g^*$

Computing the spectral density: general method

- $G_z(g)$ admits a unique zero and is holomorphic. The Newton-Raphson method is simple and efficient, but requires

$$\lim_{\varepsilon \to 0^+} G_{\lambda + i\varepsilon}(g) = G_\lambda(g),$$

which is garanteed by Silverstein and Choi, 1995.

- **Require:** λ, $k_{\max}$, tol

  $g_0 \leftarrow$ random starter in $\mathbb{C} \setminus \mathbb{R}$
  for $k = 1$ to $k_{\max}$ **do**
  $\quad g_k \leftarrow g_{k-1} - G_\lambda(g_{k-1}) / G'_\lambda(g_{k-1})$
  $\quad$**if** $|g_k - g_{k-1}| \leq$ tol **then**
  $\quad\quad g^* \leftarrow g_k$
  $\quad\quad$**exit loop**
  $\quad$**end if**
  end for
  $p(\lambda) \leftarrow \frac{1}{\pi} \text{Im } g^*$

Spectral density of POW1 estimators

- Closed-form results outperform the purely numerical approach. For $\alpha(t) \sim 1/(1+(\frac{t}{c})^{\gamma})$, calculations in the limit $T/N = \infty$ lead to

$$1 + zg(z) = g(z)(1 + K\,g(z))^{\frac{1}{\gamma}-1} \quad \gamma > 1.$$

- Writing $\gamma = q/p$, with $q > p \geq 1$ two integers yields

$$(1 + Kg(z))^{q-p}(1 + zg(z))^{q} - g(z)^{q} = 0.$$

POW1 spectral density has an explicit form only for $\gamma = 2$ and $\gamma = 3/2$.

- Exact calculations sometimes possible when $T/N = c_0 < \infty$. E.g. $\gamma = 1$:

$$1 + zg(z) = cKg(z)\log\left(1 + \frac{c_0}{c(1 + Kg(z))}\right).$$

Expression for $\gamma = 1/p$, $p \in \mathbb{N}$ in (M. de Lachapelle, Lévèque, 2009).

Spectral density of POW1 estimators

- Closed-form results outperform the purely numerical approach. For $\alpha(t) \sim 1/(1+(\frac{t}{c})^\gamma)$, calculations in the limit $T/N = \infty$ lead to

$$1 + zg(z) = g(z)(1 + K\,g(z))^{\frac{1}{\gamma}-1} \quad \gamma > 1.$$

- Writing $\gamma = q/p$, with $q > p \geq 1$ two integers yields

$$(1 + Kg(z))^{q-p}(1 + zg(z))^q - g(z)^q = 0.$$

  POW1 spectral density has an explicit form only for $\gamma = 2$ and $\gamma = 3/2$.

- Exact calculations sometimes possible when $T/N = c_0 < \infty$. E.g. $\gamma = 1$:

$$1 + zg(z) = cKg(z)\log\left(1 + \frac{c_0}{c(1+Kg(z))}\right).$$

  Expression for $\gamma = 1/p$, $p \in \mathbb{N}$ in (M. de Lachapelle, Lévèque, 2009)

Spectral density of POW1 estimators

- Closed-form results outperform the purely numerical approach. For $\alpha(t) \sim 1/(1+(\frac{t}{c})^{\gamma})$, calculations in the limit $T/N = \infty$ lead to

$$1 + zg(z) = g(z)(1 + K\,g(z))^{\frac{1}{\gamma}-1} \quad \gamma > 1.$$

- Writing $\gamma = q/p$, with $q > p \geq 1$ two integers yields

$$(1 + Kg(z))^{q-p}(1 + zg(z))^{q} - g(z)^{q} = 0.$$

  POW1 spectral density has an explicit form only for $\gamma = 2$ and $\gamma = 3/2$.

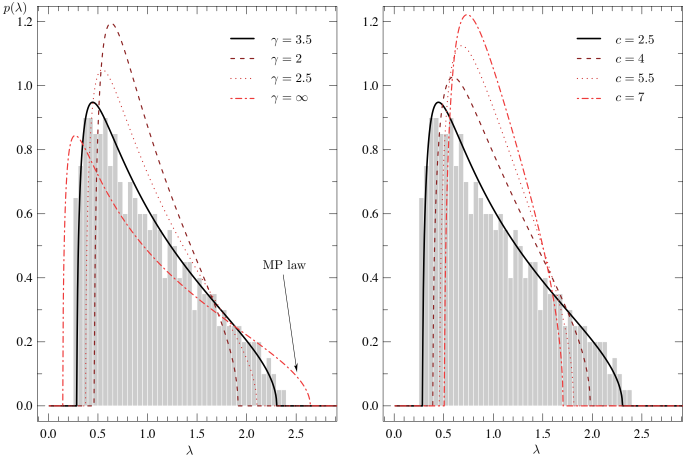
- Exact calculations sometimes possible when $T/N = c_0 < \infty$.
  E.g. $\gamma = 1$:

$$1 + zg(z) = cKg(z)\log\left(1 + \frac{c_0}{c\,(1 + Kg(z))}\right).$$

  Expression for $\gamma = 1/p$, $p \in \mathbb{N}$ in (M. de Lachapelle, Lévèque, 2009).

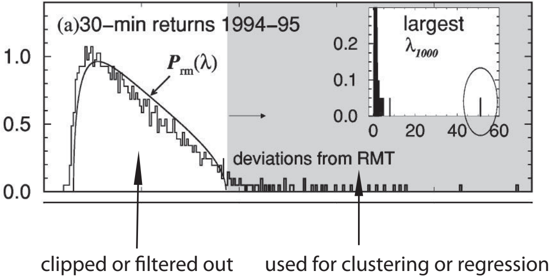
Spectral density of POW1 estimators: plots

- Spectral histogram of a 400×2000 correlation matrix of i.i.d Student returns and asymptotic density in the limit $T/N \to \infty$.



Applications

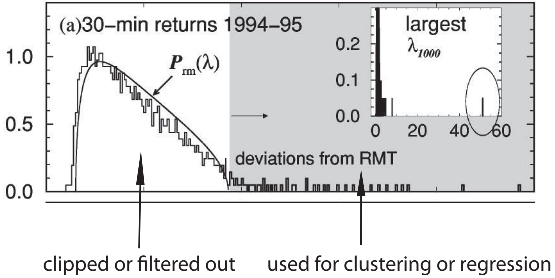
- In portfolio allocation, Random Matrix Theory (RMT) is used to locate informative eigenpairs (λ_i, v_i).



- RMT can also be used to control the effects of the weighting on the matrix conditioning $\lambda_{max}/\lambda_{min}$ and "noise band" $\lambda_{max} - \lambda_{min}$.

Applications

- In portfolio allocation, Random Matrix Theory (RMT) is used to locate informative eigenpairs (λ_i, v_i).



- RMT can also be used to control the effects of the weighting on the matrix conditioning $\lambda_{\max}/\lambda_{\min}$ and "noise band" $\lambda_{\max} - \lambda_{\min}$.

Edge spectrum of POW estimators

- Marčenko and Pastur, 1967: The frontiers of the spectrum are extrema of g^{-1}. Convenient, since from the expression of the Stieltjes transform, g^{-1} is always known explicitely. For POW1, it reads

$$g^{-1}(y) = (1+Ky)^{\frac{1}{\gamma}-1} - \frac{1}{y}.$$

  The edge spectrum is defined as $g^{-1}(y_{\pm})$, where $y_{\pm}$ are the only solutions to $(g^{-1}(y_{\pm}))' = 0$. For POW1:

$$\lambda_{1,N} = \left(\frac{K \mp \sqrt{K(K+4(\gamma-1)\gamma\lambda_{1,N}}}{2(\gamma-1)\lambda_{1,N}} + 1\right)^{\frac{1}{\gamma}-1} \pm \frac{2(\gamma-1)\lambda_{1,N}}{\sqrt{\frac{4(\gamma-1)\gamma\lambda_{1,N}}{K}+1} \mp 1}.$$

- The critical regime is $\lambda_N \to 0^+$ (conditioning issues). Expanding to first order in λ_N and solving leads to

$$\lambda_N \underset{\lambda_N \to 0^+}{\approx} c^{\gamma-1}$$

Edge spectrum of POW estimators

- Marčenko and Pastur, 1967: The frontiers of the spectrum are extrema of g^{-1}. Convenient, since from the expression of the Stieltjes transform, g^{-1} is always known explicitely. For POW1, it reads

$$g^{-1}(y) = (1+Ky)^{\frac{1}{\gamma}-1} - \frac{1}{y}.$$

The edge spectrum is defined as $g^{-1}(y_\pm)$, where $y_\pm$ are the only solutions to $(g^{-1}(y_\pm))' = 0$. For POW1:

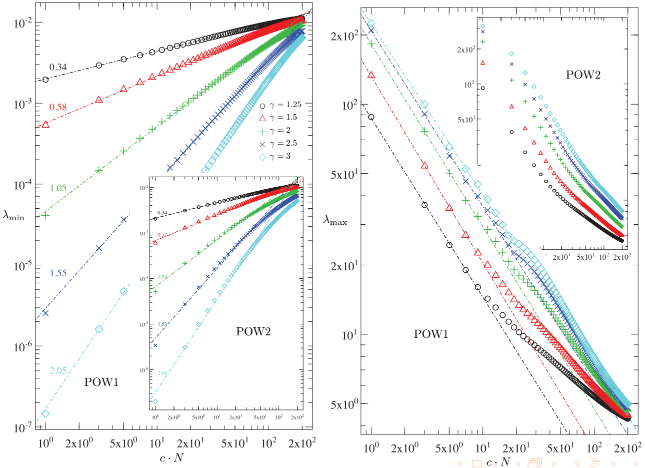
$$\lambda_{1,N} = \left(\frac{K \mp \sqrt{K(K+4(\gamma-1)\gamma\lambda_{1,N}}}{2(\gamma-1)\lambda_{1,N}} + 1 \right)^{\frac{1}{\gamma}-1} \pm \frac{2(\gamma-1)\lambda_{1,N}}{\sqrt{\frac{4(\gamma-1)\gamma\lambda_{1,N}}{K}+1} \mp 1}.$$

- The critical regime is $\lambda_N \to 0^+$ (conditioning issues). Expanding to first order in λ_N and solving leads to

$$\lambda_N \underset{\lambda_N \to 0^+}{\approx} c^{\gamma-1}$$

Edge spectrum of POW estimators: Plots

- Edge spectrum of $H \in \mathbb{R}^{400 \times 500}$, where 400 independent Student returns with mean zero and degree of freedom chosen at random in $[2, 5]$.



Value-at-Risk (VaR) estimation

- The VaR represents the maximum loss associated with this position during the holding period for a given confidence level probability.
- It is defined as

$$p = P(\Delta V_{t,h} \leq \mathrm{VaR}_{t,h}(p)),$$

where $\Delta V_{t,h} = \mathrm{price}(t+h) - \mathrm{price}(t)$.

- Here VaR is estimated historically with the three profiles.
- The quality of the estimation is compared with realized VaR violations

EPFL
SWISSQUOTE

Value-at-Risk (VaR) estimation

- The VaR represents the maximum loss associated with this position during the holding period for a given confidence level probability.
- It is defined as
$$p = P(\Delta V_{t,h} \leq \mathrm{VaR}_{t,h}(p)),$$
where $\Delta V_{t,h} = \mathrm{price}(t+h) - \mathrm{price}(t)$.
- Here VaR is estimated historically with the three profiles.
- The quality of the estimation is compared with realized VaR violations

Value-at-Risk (VaR) estimation

- The VaR represents the maximum loss associated with this position during the holding period for a given confidence level probability.
- It is defined as

$$p = P(\Delta V_{t,h} \leq \mathrm{VaR}_{t,h}(p)),$$

  where $\Delta V_{t,h} = \mathrm{price}(t+h) - \mathrm{price}(t)$.
- Here VaR is estimated historically with the three profiles.
- The quality of the estimation is compared with realized VaR violations

Value-at-Risk (VaR) estimation

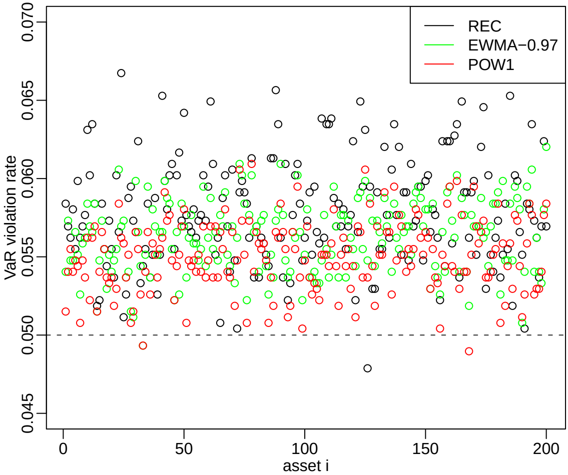
- The VaR represents the maximum loss associated with this position during the holding period for a given confidence level probability.
- It is defined as

$$p = P(\Delta V_{t,h} \leq \mathrm{VaR}_{t,h}(p)),$$

where $\Delta V_{t,h} = \mathrm{price}(t+h) - \mathrm{price}(t)$.

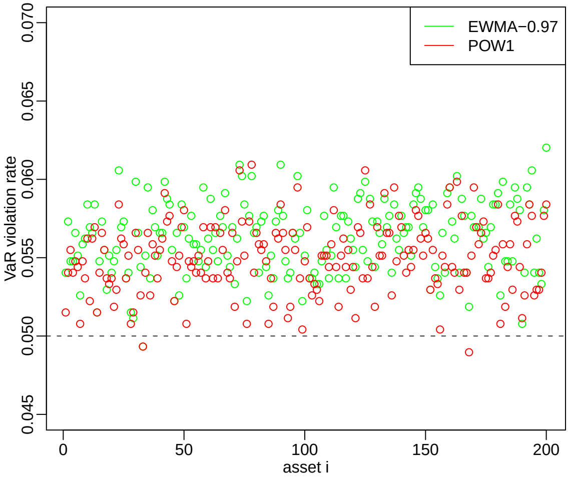
- Here VaR is estimated historically with the three profiles.
- The quality of the estimation is compared with realized VaR violations

Value-at-Risk (VaR) estimation: results



Value-at-Risk (VaR) estimation: results

- EWMA-0.97 -> 33 mean nb of days. POW1 parameters are $\gamma = 2$ and c so as to have 33 days. No optimization (not yet!).



Conclusions

- Weighted (co)variance matrices induce stochastic processes that can be analysed by econometric techniques. Results show return excess Kurtosis and volatility long-range autocorrelation (depending on weighting).
- Power-law decaying weights "spare" the edge spectrum of weighted correlation matrices while still doing good at capturing volatility risks.
- Bunch of applications in financial risk assessment and portfolio allocation.

Conclusions

- Weighted (co)variance matrices induce stochastic processes that can be analysed by econometric techniques. Results show return excess Kurtosis and volatility long-range autocorrelation (depending on weighting).
- Power-law decaying weights "spare" the edge spectrum of weighted correlation matrices while still doing good at capturing volatility risks.
- Bunch of applications in financial risk assessment and portfolio allocation.

EPFL SWISSQUOTE

Conclusions

- Weighted (co)variance matrices induce stochastic processes that can be analysed by econometric techniques. Results show return excess Kurtosis and volatility long-range autocorrelation (depending on weighting).
- Power-law decaying weights "spare" the edge spectrum of weighted correlation matrices while still doing good at capturing volatility risks.
- Bunch of applications in financial risk assessment and portfolio allocation.

Acknowledgement

- The Quantitative Asset Management team at Swissquote.
- Our peers, Vladimir Marčenko and Leonid Pastur