# Applications of Free Random Variables to Financial Analysis

Andrzej Jarosz andrzej.jarosz@clico.pl Clico Ltd., Kraków, Poland

in collaboration with Z. Burda, J. Jurkiewicz, M. A. Nowak, M. Snarska (Jagiellonian University, Kraków, Poland), G. Papp (Eötvös University, Budapest, Hungary), I. Zahed (SUNY Stony Brook, USA)

## Abstract

We apply the concept of *free random variables* (FRV), which is a noncommutative extension of probability calculus, to *doubly–correlated Gaussian Wishart* random matrix models, appearing for example in multivariate analysis of financial time series displaying both inter–asset cross–covariances $\mathbf{C}$ (such as in factor models) and temporal auto–covariances $\mathbf{A}$ (such as in models with heteroscedasticity, in VARMA processes, *etc.*).

## Correlated Gaussians

**We consider** a universe of $N$ financial assets ($i = 1, \ldots, N \to \infty$), sampled over $T$ time moments ($a = 1, \ldots, T \to \infty$): $R_{ia}$ is *e.g.* the demeaned logarithmic return. The simplest approximations: $\mathbf{R}$ is a *Gaussian* random matrix with the structure of covariances,

$$\langle R_{ia} R_{jb} \rangle = C_{ij} A_{ab}. \qquad (1)$$

The change of variables $\mathbf{R} = \sqrt{\mathbf{C}}\widetilde{\mathbf{R}}\sqrt{\mathbf{A}}$ gives uncorrelated Gaussians $\widetilde{\mathbf{R}}$.

**Problem:** *estimation* of $\mathbf{C}$ from historical time series. Marred by the measurement *noise*, quantified by $r \equiv N/T$. The Pearson estimator of $\mathbf{C}$,

$$\mathbf{c} = \frac{1}{T}\sqrt{\mathbf{C}}\widetilde{\mathbf{R}}\mathbf{A}\widetilde{\mathbf{R}}^{\mathrm{T}}\sqrt{\mathbf{C}}. \qquad (2)$$

## Cross– and Auto–Covariances

**Models of cross–covariances:** Typical spectra of $\mathbf{c}$: a few large eigenvalues ("factors") plus a "sea" of the Marčenko–Pastur distribution. The largest factor represents the market, the other factors — the industrial sectors.

**Models of auto–covariances:**

- *Heteroscedasticity* (volatility clustering) is a phenomenon of a stochastic time dependence of volatility $\sigma_{ia}$. These lagged correlations of volatility are not strong (a few percent), but persist over long periods of time. Typically modeled by a power–law decay, $\langle \sigma_a^2 \sigma_b^2 \rangle - \langle \sigma_a^2 \rangle \langle \sigma_b^2 \rangle \propto 1/|b-a|^{\nu}$ ($\nu \sim 0.2 \div 0.4$). An I–GARCH(1) (EWMA) model suggests a less realistic exponential decay $\mathrm{e}^{-|b-a|/\tau}$, yet widely exploited in the financial industry (RiskMetrics 1994, $\tau = 16.2$ days).
- Exponentially decaying *lagged correlations between the residuals* $R_{ia}/\sigma_{ia}$, albeit weak and short–ranged, are included in the new RiskMetrics 2006.

Other models in other contexts (*e.g.* in wireless communication).

## Free Random Variables

Voiculescu and Speicher's [4] *free random variables calculus* is a generalization of probability theory to *noncommutative random variables*, such as infinite (Hermitian) random matrices $\mathbf{X}$. It relies on the concept of *freeness*, which is noncommutative independence.

| **Classical probability:** | **Noncommutative probability (FRV):** |
|---|---|
| p.d.f., $P_X(x)$ | spectral density, $\rho_{\mathbf{X}}(\lambda)$ |
| characteristic function, $g_X(z) \equiv \langle \mathrm{e}^{\mathrm{i}zX} \rangle$ | Green's function, $G_{\mathbf{X}}(z) \equiv (1/N)\langle \mathrm{Tr} 1/(z\mathbf{1}_N - \mathbf{X}) \rangle$, or $M$–transform, $M_{\mathbf{X}}(z) = zG_{\mathbf{X}}(z) - 1$ |
| independence | freeness |
| **Addition of independent commutative random variables:** The logarithm of the characteristic function, $r_X(z) \equiv \log g_X(z)$, is additive, $$r_{X_1+X_2}(z) = r_{X_1}(z) + r_{X_2}(z). \qquad (3)$$ | **Addition of free noncommutative random variables:** The Blue's function, $G_{\mathbf{X}}(B_{\mathbf{X}}(z)) = B_{\mathbf{X}}(G_{\mathbf{X}}(z)) = z$, is additive, $$B_{\mathbf{X}_1+\mathbf{X}_2}(z) = B_{\mathbf{X}_1}(z) + B_{\mathbf{X}_2}(z) - \frac{1}{z}. \qquad (4)$$ |
| **Multiplication of independent r.v.:** Reduced to the addition problem via the exponential map, owing to $\mathrm{e}^{X_1}\mathrm{e}^{X_2} = \mathrm{e}^{X_1+X_2}$. | **Multiplication of free r.v.:** The $N$–transform, $M_{\mathbf{X}}(N_{\mathbf{X}}(z)) = N_{\mathbf{X}}(M_{\mathbf{X}}(z)) = z$, is multiplicative, $$N_{\mathbf{X}_1\mathbf{X}_2}(z) = \frac{z}{1+z} N_{\mathbf{X}_1}(z) N_{\mathbf{X}_2}(z). \qquad (5)$$ |

## The Main Result

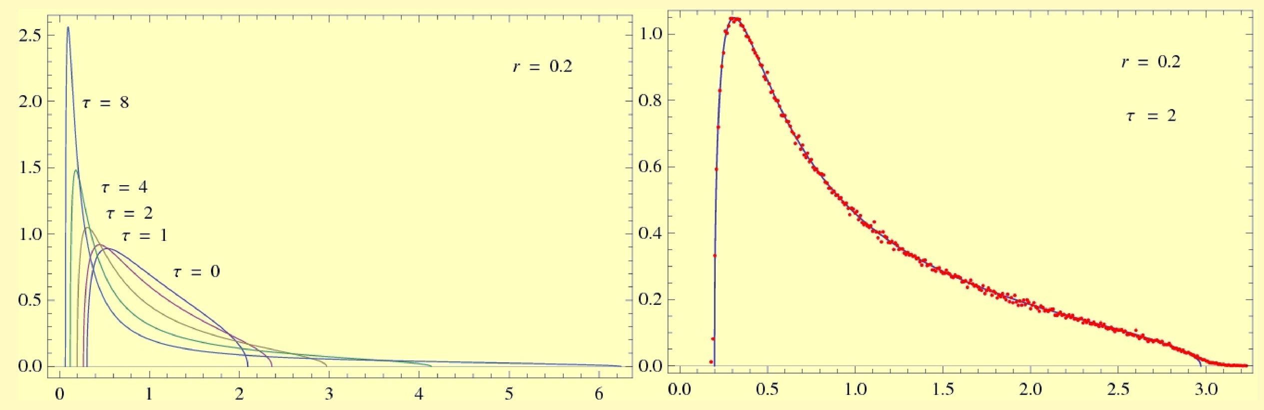
The FRV multiplication formula (5) leads via a back–of–an–envelope calculation to an equation for the $M$–transform of the Pearson estimator $\mathbf{c}$ (2) [1],

$$z = rMN_{\mathbf{A}}(rM)N_{\mathbf{C}}(M). \qquad (6)$$

## Examples

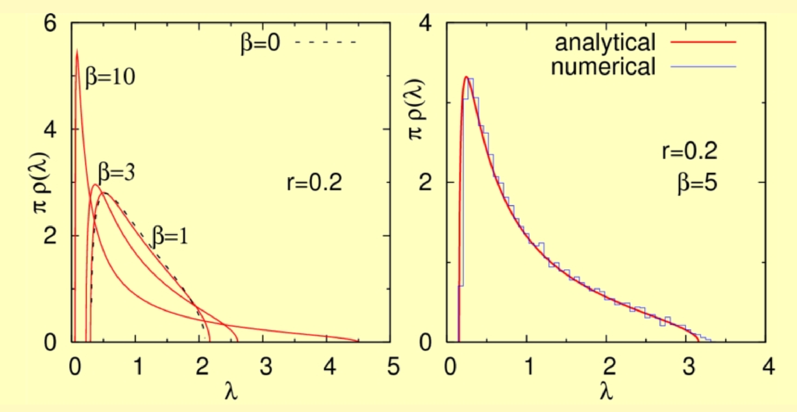
**Exponentially decaying auto–covariances:** $A_{ab} = \mathrm{e}^{-|a-b|/\tau}$; denote $\chi \equiv \coth(1/\tau)$. The main FRV formula (6) yields a 4–th order polynomial equation for $M \equiv M_{\mathbf{c}}(z)$ [1],

$$r^2M^4 + 2r(r - \chi z)M^3 + \left(z^2 - 2r\chi z + r^2 - 1\right)M^2 - 2M - 1 = 0. \qquad (7)$$



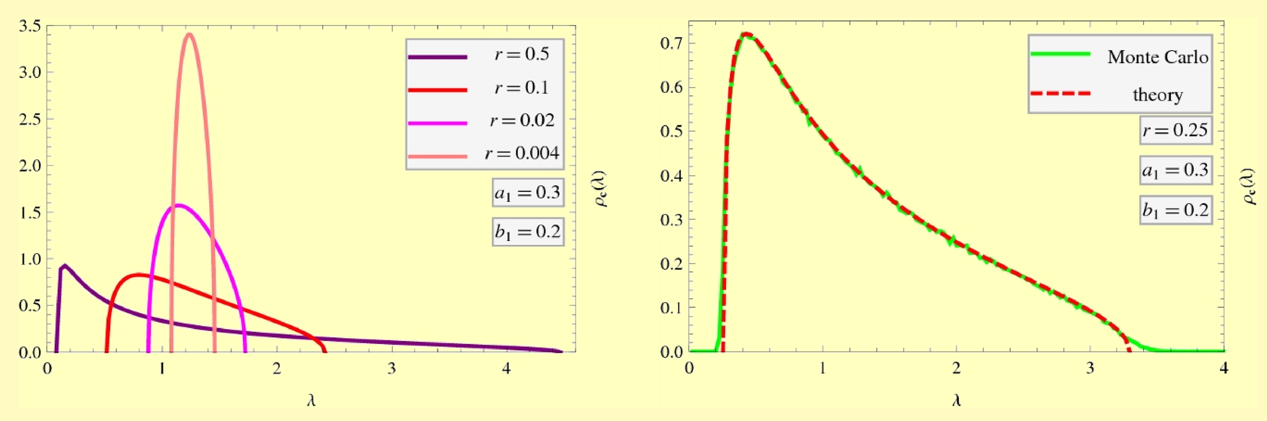
**Exponentially weighted moving average (EWMA):** $A_{ab} \equiv T\frac{1-\alpha}{1-\alpha^T}\alpha^{a-1}\delta_{ab}$, in the limit $\alpha \to 1^-$, such that $\beta \equiv T(1-\alpha)$ finite (in RiskMetrics 1994, $\alpha = 0.94$). The main FRV formula (6) yields an entangled equation for $M \equiv M_{\mathbf{c}}(z)$ [1],

$$z = \frac{r\beta(1+M)\left(\mathrm{e}^{\beta(1+rM)} - 1\right)}{\left(\mathrm{e}^{\beta} - 1\right)\left(\mathrm{e}^{\beta rM} - 1\right)}. \qquad (8)$$




**VARMA(1,1):** $R_{ia} - b_1 R_{i,a-1} = a_0\epsilon_{ia} + a_1\epsilon_{i,a-1}$ (a stochastic process), with $\epsilon_{ia}$ white noise. Its auto–covariance matrix $\mathbf{A}$ is an exponential decay with an additional constant term on the diagonal. The main FRV formula (6) yields a 6–th order polynomial equation for $M \equiv M_{\mathbf{c}}(z)$ [2],

$$r^4 a_0^2 a_1^2 \left(a_0^2 - a_1^2\right)^2 M^6 + \ldots = 0. \qquad (9)$$




Everywhere $\mathbf{C} = \mathbf{1}_N$, *i.e.*, trivial cross–covariances. The graphs show the mean spectral density $\rho_{\mathbf{c}}(\lambda)$ inferred from FRV equations (7), (8), (9) (left), and their Monte Carlo numerical confirmation (right).

## Prospects

Incorporation of non–Gaussian effects:

- Free Lévy (stable w.r.t. FRV) vs. standard Lévy [3].
- Stochastic volatility models (modeling of heteroscedasticity): $\mathbf{C}$ and $\mathbf{A}$ random instead of constant ("random deformations"); our main formula (6) remains usable.

Inclusion of a nontrivial cross–correlation structure $\mathbf{C}$ inferred from factor models. Problem: Large ($\sim N$) eigenvalues (the market and industrial sectors) of $\mathbf{C}$ not compatible with FRV.

More general correlation structure than just $\mathbf{C}$ and $\mathbf{A}$ (1). Goal: Taking into account the Epps effect.

## References

[1] Burda Z., Jarosz A., Jurkiewicz J., Nowak M. A., Papp G., Zahed I., *Applying Free Random Variables to Random Matrix Analysis of Financial Data*, submitted to Quantitative Finance.

[2] Burda Z., Jarosz A., Nowak M. A., Snarska M., *A Free Random Variables Approach to the VARMA Processes*, submitted to New J. Phys.

[3] Burda Z., Jarosz A., Jurkiewicz J., Nowak M. A., Papp G., Zahed I., *Free Lévy–Wishart Random Matrices with Spatial and Temporal Correlations*, in preparation.

[4] Voiculescu D. V., Dykema K. J., Nica A., *Free Random Variables*, Am. Math. Soc., Providence, 1992. Speicher R., Math. Ann. **298** (1994) 611.

## Funding

This work has been supported by the Polish Ministry of Science Grant No. N N202 229137 (2009–2012), and (AJ) by Clico Ltd.