

Sei $X \in \mathbb{C}^{n \times m}$ eine Matrix dessen Elemente unabhängig und identisch verteilte Zufallsvariablen sind und ihr Erwartungswert 0 mit Varianz σ^2 ist. Weiter sind $x_k = (x_{1k}, \dots, x_{nk})$, $\mathbf{X} = (x_1, \dots, x_m)$ und $\bar{x} = \frac{1}{m} \sum_{k=1}^m x_k$. Dann ist die empirische Kovarianzmatrix

$$\mathbf{S} = \frac{1}{m-1} \sum_{k=1}^m (x_k - \bar{x})(x_k - \bar{x})^H$$

0.1 Marčenko-Pastur Momente

Die Dichtefunktion der Marčenko-Pastur-Verteilung $F_y(x)$ ist durch

$$p_y(x) = \begin{cases} \frac{1}{2\pi xy\sigma^2} \sqrt{(b-x)(x-a)}, & \text{wenn } a \leq x \leq b \\ 0, & \text{sonst} \end{cases} \quad (0.1)$$

gegeben. Dabei sind $a = \sigma^2 (1 - \sqrt{y})^2$ und $b = \sigma^2 (1 + \sqrt{y})^2$ jeweils mit $y = \frac{n}{m}$ die Träger der Dichtefunktion. Wenn $\sigma^2 = 1$ ist, dann sprechen wir von der standard M-P-Verteilung. Die k-ten Momente der M-P-Verteilung sind durch

$$\beta_k = \beta_k(y, \sigma^2) := \int_a^b x^k p_y(x) dx \quad (0.2)$$

definiert. Wir können die Momente der M-P-Verteilung durch ihren Standardfall beschreiben.

Lemma 0.1. Für alle $k \geq 1$ gilt

$$\beta_k(y, \sigma^2) = \sigma^{2k} \beta_k(y, 1).$$

Beweis. Nach Definition 0.2 haben wir

$$\begin{aligned} \beta_k(y, \sigma^2) &= \int_a^b x^k p_y(x) dx \\ &= \int_{\sigma^2(1-\sqrt{y})^2}^{\sigma^2(1+\sqrt{y})^2} x^k \frac{1}{2\pi xy\sigma^2} \sqrt{(\sigma^2(1+\sqrt{y})^2 - x)(x - \sigma^2(1-\sqrt{y})^2)} dx. \end{aligned}$$

Substituieren wir mit $x = u\sigma^2$ erhalten wir

$$\begin{aligned}
 \beta_k(y, \sigma^2) &= \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} (u\sigma^2)^k \frac{1}{2\pi uy\sigma^2} \sqrt{\left(\sigma^2(1+\sqrt{y})^2 - u\sigma^2\right)\left(u\sigma^2 - \sigma^2(1-\sqrt{y})^2\right)} du \\
 &= \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} u^k \sigma^{2k} \frac{1}{2\pi uy\sigma^2} \sqrt{\sigma^4((1+\sqrt{y})^2 - u)(u - (1-\sqrt{y})^2)} du \\
 &= \sigma^{2k} \int_{(1-\sqrt{y})^2}^{(1+\sqrt{y})^2} u^k \frac{1}{2\pi uy} \sqrt{(1+\sqrt{y})^2 - u} \sqrt{u - (1-\sqrt{y})^2} du \\
 &= \sigma^{2k} \beta_k(y, 1).
 \end{aligned}$$

□

Damit können wir die Momente explizit bestimmen.

Lemma 0.2. Die explizite Darstellung der Momente ist

$$\beta_k = \sum_{r=0}^{k-1} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} y^r. \quad (0.3)$$

Beweis. Wir haben nach Definition 0.2

$$\beta_k = \frac{1}{2\pi y} \int_a^b x^{k-1} \sqrt{(b-x)(x-a)} dx.$$

Substituieren wir mit $x = 1 + y + z$ und setzen a und b ein, erhalten wir

$$\begin{aligned}
 \beta_k &= \frac{1}{2\pi y} \int_{-2\sqrt{y}}^{2\sqrt{y}} (1+y+z)^{k-1} \sqrt{\left((1+\sqrt{y})^2 - 1 - y - z\right)\left(1+y+z - (1-\sqrt{y})^2\right)} dz \\
 &= \frac{1}{2\pi y} \int_{-2\sqrt{y}}^{2\sqrt{y}} (1+y+z)^{k-1} \sqrt{4y - z^2} dz.
 \end{aligned}$$

Durch Einsetzen des Binomischen Lehrsatzes lässt sich das Integral zu

$$\begin{aligned}
 \beta_k &= \frac{1}{2\pi y} \int_{-2\sqrt{y}}^{2\sqrt{y}} \sum_{l=0}^{k-1} \binom{k-1}{l} (1+y)^{k-1-l} z^l \sqrt{4y - z^2} dz \\
 &= \frac{1}{2\pi y} \sum_{l=0}^{k-1} \binom{k-1}{l} (1+y)^{k-1-l} \int_{-2\sqrt{y}}^{2\sqrt{y}} z^l \sqrt{4y - z^2} dz
 \end{aligned}$$

umformen. Um das Integral weiter zu vereinfachen, substituieren wir mit $z = 2\sqrt{y}u$ und

erhalten

$$\begin{aligned}\beta_k &= \frac{1}{2\pi y} \sum_{l=0}^{k-1} \binom{k-1}{l} (1+y)^{k-1-l} \int_{-1}^1 (2\sqrt{y}u)^l \sqrt{4y - 4yu^2} du \\ &= \frac{1}{2\pi y} \sum_{l=0}^{k-1} \binom{k-1}{l} (1+y)^{k-1-l} (2\sqrt{y})^l 4y \int_{-1}^1 u^l \sqrt{1-u^2} du.\end{aligned}$$

Hier ist unser Integrand ungerade, sofern l ungerade ist. Da die Integralgrenzen symmetrisch um den Nullpunkt liegen ist das Integral für alle ungeraden l gleich 0. Dadurch ist jeder zweite Summand 0. Wir können also jeden zweiten Summanden überspringen. Dafür lassen wir die Summe bis zur Hälfte von $k-1$ laufen und summieren über $2l$ statt l . Für den Fall, dass $k-1$ ungerade ist, runden wir ab, weil es dann nur $\frac{k-1}{2}$ Summanden gibt die ungleich 0 sind. Wir haben also

$$\begin{aligned}\beta_k &= \frac{1}{2\pi y} \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \binom{k-1}{2l} (1+y)^{k-1-2l} (2\sqrt{y})^{2l} 4y \int_{-1}^1 u^{2l} \sqrt{1-u^2} dz \\ &= \frac{1}{2\pi y} \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \binom{k-1}{2l} (1+y)^{k-1-2l} (4y)^{l+1} \int_{-1}^1 u^{2l} \sqrt{1-u^2} dz.\end{aligned}$$

Hier ist der Integrand eine gerade Funktion für alle $0 \leq l$. Da die Integralgrenzen symmetrisch um den Nullpunkt liegen können wir auch zwei mal über das halbe Intervall integrieren. Folglich gilt

$$\beta_k = \frac{1}{2\pi y} \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \binom{k-1}{2l} (1+y)^{k-1-2l} (4y)^{l+1} 2 \int_0^1 u^{2l} \sqrt{1-u^2} dz.$$

Wenn wir jetzt mit $u = \sqrt{w}$ substituieren,

$$\begin{aligned}\beta_k &= \frac{1}{2\pi y} \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \binom{k-1}{2l} (1+y)^{k-1-2l} (4y)^{l+1} \int_0^1 w^l \sqrt{1-w} \frac{1}{\sqrt{w}} dw \\ &= \frac{1}{2\pi y} \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \binom{k-1}{2l} (1+y)^{k-1-2l} (4y)^{l+1} \int_0^1 w^{l-\frac{1}{2}} \sqrt{1-w} dw\end{aligned}$$

erhalten wir ein Integral, dass wir mittels Eulerschen Betafunktion [Quelle!](#) lösen können.

Die Eulersche Betafunktion ist durch

$$\int_0^1 t^{p-1} (1-t)^{q-1} dt := \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

definiert. Wir können also mit $p = l + \frac{1}{2}$ und $q = 1 + \frac{1}{2}$ das Integral durch die Gammafunktionen

$$\begin{aligned} \int_0^1 w^{l-\frac{1}{2}} \sqrt{1-w} dw &= \frac{\Gamma(l+\frac{1}{2})\Gamma(1+\frac{1}{2})}{\Gamma(l+2)} \\ &= \frac{\frac{(2l)!\sqrt{\pi}}{l!4^l} \frac{2!\sqrt{\pi}}{4^l}}{(l+1)!} \\ &= \frac{\frac{(2l)!}{l!4^l} \frac{1}{2}\pi}{(l+1)!} \end{aligned} \tag{0.4}$$

beschreiben. Nach Einsetzen von 0.4 erhalten wir

$$\begin{aligned} \beta_k &= \frac{1}{2\pi y} \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \binom{k-1}{2l} (1+y)^{k-1-2l} (4y)^{l+1} \frac{\frac{(2l)!\frac{1}{2}\pi}{l!4^l}}{(l+1)!} \\ &= \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \frac{1}{2\pi y} \frac{(k-1)!}{(2l)!(k-1-2l)!} (1+y)^{k-1-2l} (4y)^{l+1} \frac{\frac{(2l)!\frac{1}{2}\pi}{l!4^l}}{(l+1)!}. \end{aligned}$$

Hier lassen sich diverse Variablen so kürzen, dass wir

$$\beta_k = \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \frac{(k-1)!}{l!(l+1)!(k-1-2l)!} y^l (1+y)^{k-1-2l}$$

erhalten. Wenn wir den Binomischen Lehrsatz mit

$$\begin{aligned} (1+y)^{k-1-2l} &= \sum_{s=0}^{k-1-2l} \binom{k-1-2l}{s} y^s \\ &= \sum_{s=0}^{k-1-2l} \frac{(k-1-2l)!}{s!(k-1-2l-s)!} y^s \end{aligned}$$

einsetzen, haben wir

$$\begin{aligned}\beta_k &= \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \sum_{s=0}^{k-1-2l} \frac{(k-1)!}{l!(l+1)!(k-1-2l)!} y^l \frac{(k-1-2l)!}{s!(k-1-2l-s)!} y^s \\ &= \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \sum_{s=0}^{k-1-2l} \frac{(k-1)!}{l!(l+1)!s!(k-1-2l-s)!} y^{l+s}.\end{aligned}$$

Jetzt wollen wir den Term wieder vereinfachen. Dafür setzen wir $r = s + l$

$$\beta_k = \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \sum_{s=l}^{k-1-l} \frac{(k-1)!}{l!(l+1)!(r-l)!(k-1-2l-s)!} y^r$$

und erweitern mit zwei Einsen

$$\begin{aligned}\beta_k &= \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \sum_{s=l}^{k-1-l} \frac{(k-1)!}{l!(l+1)!(r-l)!(k-1-2l-s)!} \frac{r!}{r!} \frac{(k-r)!}{(k-r)!} y^r \\ &= \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \sum_{s=l}^{k-1-l} \frac{(k-1)!}{r!(k-r)!} \frac{r!}{l!(l-r)!} \frac{(k-r)!}{(k-r-l-1)!(l+1)!} y^r \\ &= \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \sum_{s=l}^{k-1-l} \frac{1}{k} \frac{k!}{r!(k-r)!} \frac{r!}{l!(l-r)!} \frac{(k-r)!}{(l+1)!(k-r-(l+1))!} y^r \\ &= \sum_{l=0}^{\lfloor \frac{k-1}{2} \rfloor} \sum_{s=l}^{k-1-l} \frac{1}{k} \binom{k}{r} \binom{r}{l} \binom{k-r}{l+1} y^r.\end{aligned}$$

Für den nächsten Schritt müssen wir die Laufvariablen voneinander unabhängig machen.

Momentan haben wir erst das l und ermitteln damit, inwieweit das r eingeschränkt ist.

Für ein festes aber beliebiges k liegen unsere Laufvariablen in

$$A = \left\{ (l, r) \in \mathbb{N}^2 \mid 0 \leq l \leq \left\lfloor \frac{k-1}{2} \right\rfloor, l \leq r \leq k-1-l \right\}.$$

Daraus wissen wir, dass

$$0 \leq l, r \leq k-1-l$$

gilt, woraus wir

$$l \leq k-1-r$$

folgern können. Damit erhalten wir

$$\begin{aligned} 0 &\leq l \leq r \leq k-1 - l \leq k-1 \\ \implies 0 &\leq r \leq k-1, 0 \leq l \leq \min(r, k-1-r) \end{aligned}$$

damit können wir also die Summen äquivalent über

$$A = \{(l, r) \in \mathbb{N}^2 \mid 0 \leq r \leq k-1, 0 \leq l \leq \min(r, k-1-r)\} \quad (0.5)$$

laufen lassen.

Wir haben also mit 0.5

$$\begin{aligned} \beta_k &= \sum_{r=0}^{k-1} \sum_{l=0}^{\min(r, k-1-r)} \frac{1}{k} \binom{k}{r} \binom{r}{l} \binom{k-r}{l+1} y^r \\ &= \sum_{r=0}^{k-1} \frac{1}{k} \binom{k}{r} y^r \sum_{l=0}^{\min(r, k-1-r)} \binom{r}{l} \binom{k-r}{l+1}. \end{aligned}$$

Das Minimum können wir noch weiter vereinfachen, da die Summanden gleich 0 sind, sobald $l > r$ oder $l > k-1-r$ ist. Wenn wir die Summe also bis r laufen lassen, ändern wir nichts am Ergebnis.

$$\beta_k = \sum_{r=0}^{k-1} \frac{1}{k} \binom{k}{r} y^r \sum_{l=0}^r \binom{r}{l} \binom{k-r}{l+1}.$$

An diesem Punkt müssen wir nur noch die innere Summe vereinfachen. Dafür machen wir einen Indexshift und formen die Binomialkoeffizienten etwas um. Wir haben also

$$\begin{aligned} \beta_k &= \sum_{r=0}^{k-1} \frac{1}{k} \binom{k}{r} y^r \sum_{l=1}^{r+1} \binom{r}{l-1} \binom{k-r}{l} \\ &= \sum_{r=0}^{k-1} \frac{1}{k} \binom{k}{r} y^r \sum_{l=1}^{r+1} \binom{r}{r-(l-1)} \binom{k-r}{l} \\ &= \sum_{r=0}^{k-1} \frac{1}{k} \binom{k}{r} y^r \sum_{l=1}^{r+1} \binom{r}{(r+1)-l} \binom{k-r}{l}. \end{aligned}$$

Wir wollen hier die Vandermonde Identität [Quelle!](#)

$$\sum_{\ddot{o}=0}^{\ddot{u}} \binom{n}{\ddot{o}} \binom{m}{\ddot{u} - \ddot{o}} = \binom{n+m}{\ddot{u}}$$

nutzen. Dafür lassen wir die Summe von $l = 0$ beginnen, da der Summand das Ergebnis nicht ändert. Mit der Identität ?? erhalten wir

$$\begin{aligned}\beta_k &= \sum_{r=0}^{k-1} \frac{1}{k} \binom{k}{r} y^r \sum_{l=0}^{r+1} \binom{r}{(r+1)-l} \binom{k-r}{l} \\ &= \sum_{r=0}^{k-1} \frac{1}{k} \binom{k}{r} \binom{k}{r+1} y^r \\ &= \sum_{r=0}^{k-1} \frac{1}{k} \frac{r+1}{r+1} \binom{k}{r} \binom{k}{r+1} y^r \\ &= \sum_{r=0}^{k-1} \frac{1}{r+1} \binom{k}{r} \binom{k-1}{r} y^r.\end{aligned}$$

□