Final project for the course 18.338

Poorya Habibzadeh

December 2020

1 Deriving a closed form for the Cauchy transform of two laws

Before attempting to derive the Cauchy transforms, let us review some properties of this transformation.

Proposition 1.1. *$\Im(z)$ and $\Im(G(z))$ have the opposite sign.*

Proof. Let $z = a + ib$, we have

$$\Im\big(G(a+ib)\big) = \int_{\mathbb{R}} \frac{-f(t)b}{(t-a)^2+b^2} dt,$$

where $f(t) \geq 0$. Thus the sign of $\Im(z) = b$ and $\Im\big(G(z)\big)$ are the opposite of each other. □

Lemma 1.2. *The Cauchy transform is analytic on its domain; namely, the upper and lower half complex plane.*

Proof. Due to symmetry of the transformation, we need to prove it only for one of the half planes, say the upper half complex plane $z \in \mathbb{C}^+$. Let $\gamma \subset \mathbb{C}^+$ be any closed, piecewise differentiable curve, then

$$\oint_{\gamma} \int_{\mathbb{R}} \frac{f(t)}{z-t} dt dz = \int_{\mathbb{R}} \oint_{\gamma} \frac{f(t)}{z-t} dz dt = 0,$$

where the first equality follows by Fubini's theorem, and the second one follows by the fact that $f(t)/(z-t)$ is analytic in the upper-half complex plane. Then by Morera's theorem, the Cauchy transform must be analytic. □

Lemma 1.3. *Given the compactly supported distribution $f(x)$ with Supp$\big(f(x)\big) \subset [-r, r]$, for some $r > 0$, we have*

$$G(z) = \sum_{n=0}^{\infty} \frac{m_n}{z^{n+1}}, \qquad \text{for } |z| > r.$$

1.1 Cauchy transform of the semicircle law

Lemma 1.4. *We have the following recurrence equation for the Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$:*

$$C_n = \sum_{i=0}^{n-1} C_i C_{n-i-1}, \quad C_0 = 1, \qquad n \geq 1. \tag{1.1}$$

Theorem 1.5. *The Cauchy transform of the semicircle law is*

$$G(z) = \frac{z - \sqrt{z^2-4}}{2}. \tag{1.2}$$

Proof. We know that

$$G(z) = \sum_{k=0}^{\infty} \frac{C_k}{z^{2k+1}}, \tag{1.3}$$

for $|z| > 2$. Now, we show that it satisfies the equation $G(z)^2 - zG(z) + 1 = 0$. We have

$$\begin{aligned} G(z)^2 &= \big(C_0 z^{-1} + C_1 z^{-3} + C_2 z^{-5} + \dots\big)\big(C_0 z^{-1} + C_1 z^{-3} + C_2 z^{-5} + \dots\big) \\ &= C_0 C_0 z^{-2} + \big(C_0 C_1 + C_1 C_0\big) z^{-4} + \big(C_0 C_2 + C_1 C_1 + C_2 C_0\big) z^{-6} + \dots \\ &= \sum_{k=1}^{\infty} z^{-2k} \sum_{i=0}^{k-1} C_i C_{k-i-1} \\ &= \sum_{k=1}^{\infty} z^{-2k} C_k \\ &= z \sum_{k=0}^{\infty} z^{-2k-1} C_k - 1 \\ &= zG(z) - 1, \end{aligned} \tag{1.4}$$

where the fourth equality follows by lemma 1.4. This equation has two solutions, and only the following satisfies the condition of the Cauchy transform in Prop. 1.1:

$$\frac{z - \sqrt{z^2 - 4}}{2}.$$

This is the Cauchy transform of the semicircle for $|z| > 2$. Since it is analytic on $\mathbb{C} \setminus \mathbb{R}$, it is the Cauchy transform of the semicircle over the whole domain. ☐

Remark 1.6. Note that the Laurent series of function 1.2 at infinity is exactly the series 1.3. On the other hand, the Laurent series of the other root of the equation is

$$\frac{z + \sqrt{z^2 - 4}}{2} = z - \frac{1}{z} - \frac{1}{z^3} - \frac{2}{z^5} + O\Big(\frac{1}{z^7}\Big).$$

Thus it definitely cannot be the Cauchy transform.

1.2 Cauchy transform of the Marĉenko-Pastur law

Lemma 1.7 ([Petersen, 2015])**.** *We have the following recurrence equation for the Narayana polynomials* $N_k(r) = \sum_{j=1}^{k} N_{k,j} r^j$*, where* $N_{k,j} = \frac{1}{k}\binom{k}{j}\binom{k}{j-1}$*:*

$$N_k(r) = N_{k-1}(r) + \sum_{i=0}^{k-2} N_i(r) N_{k-i-1}(r), \quad N_0(r) = N_1(r) = r, \qquad k \geq 2. \tag{1.5}$$

Corollary 1.8. *Using the previous lemma, we can show that for* $k \geq 2$*,*

$$\sum_{i=0}^{k-1} N_i(r) N_{k-i-1}(r) = N_k(r) + (r-1) N_{k-1}(r). \tag{1.6}$$

Theorem 1.9. *The Cauchy transform of the Marĉenko-Pastur formula is*

$$G(z) = \frac{z + r - 1 - \sqrt{(z + r - 1)^2 - 4zr}}{2zr}, \tag{1.7}$$

where $r > 0$.

Proof. We need to show that

$$G(z) = \frac{1}{r}\sum_{k=0}^{\infty}\frac{N_k(r)}{z^{k+1}}, \qquad \text{for} \quad |z| > (1+\sqrt{r})^2, \tag{1.8}$$

satisfies the equation $zrG(z)^2 - (z+r-1)G(z) + 1 = 0$. We have

$$\begin{aligned}
z^2r^2G(z)^2 &= \big(N_0(r)z^0 + N_1(r)z^{-1} + N_2(r)z^{-2} + \dots\big)\big(N_0(r)z^0 + N_1(r)z^{-1} + N_2(r)z^{-2} + \dots\big)\\
&= N_0(r)N_0(r)z^0 + \big(N_0(r)N_1(r) + N_1(r)N_0(r)\big)z^{-1} + \big(N_0(r)N_2(r) + N_1(r)N_1(r) + N_2(r)N_0(r)\big)z^{-2} + \dots\\
&= \sum_{k=1}^{\infty} z^{-k+1}\sum_{i=0}^{k-1} N_i(r)N_{k-i-1}(r)\\
&= r^2 + \sum_{k=2}^{\infty} z^{-k+1}\big(N_k(r) + (r-1)N_{k-1}(r)\big)\\
&= r^2 - r + \sum_{k=1}^{\infty} z^{-k+1}N_k(r) + (r-1)\sum_{k=1}^{\infty} z^{-k}N_k(r)\\
&= \sum_{k=0}^{\infty} z^{-k+1}N_k(r) + (r-1)\sum_{k=0}^{\infty} z^{-k}N_k(r) - zr\\
&= (z^2 + zr - z)\sum_{k=0}^{\infty}\frac{N_k(r)}{z^{k+1}} - zr\\
&= zr(z+r-1)G(z) - zr,
\end{aligned} \tag{1.9}$$

where the fourth equality follows by corollary 1.8. We choose the solution that satisfies the condition of the Cauchy transform in Prop. 1.1:

$$G(z) = \frac{z+r-1-\sqrt{(z+r-1)^2 - 4zr}}{2zr}, \tag{1.10}$$

This is the Cauchy transform for $|z| > (1+\sqrt{r})^2$. Since it is analytic on $\mathbb{C}\setminus\mathbb{R}$, it is the Cauchy transform of the Marĉenko-Pastur formula over the whole domain. $\square$

Remark 1.10. Corollary 1.8 shows that if we let $r = 1$, we have $G(z) = 1/2 - \sqrt{1/4 - 1/z}$, which has the following Laurent series at infinity:

$$\frac{1}{z} + \frac{1}{z^2} + \frac{2}{z^3} + \frac{5}{z^4} + O\Big(\frac{1}{z^5}\Big),$$

while the other solution of the quadratic equation has the following series

$$G(z) = \frac{z+\sqrt{z^2-4z}}{2z} = 1 - \frac{1}{z} - \frac{1}{z^2} - \frac{2}{z^3} - \frac{5}{z^4} + O\Big(\frac{1}{z^5}\Big).$$

Therefore, we pick the solution with the minus sign.

1.3 The ratio of two consecutive Narayana polynomials:

We want to compute

$$L = \lim_{k\to\infty}\frac{N_{k+1}(r)}{N_k(r)}.$$

For $k \geq 2$ we have the following recurrence [Sulanke, 2002]:

$$(k+1)N_k(r) = (2k-1)(1+r)N_{k-1}(r) - (k-2)(r-1)^2 N_{k-2}(r). \tag{1.11}$$

Then

$$\frac{N_k(r)}{N_{k-1}(r)} = \frac{(2k-1)(1+r)}{k+1} - \frac{(k-2)(r-1)^2}{k+1}\frac{N_{k-2}(r)}{N_{k-1}(r)}. \tag{1.12}$$

This shows that the ratio $\frac{N_k(r)}{N_{k-1}(r)}$ is non-decreasing for $k \geq 2$. Moreover

$$\frac{N_k(r)}{N_{k-1}(r)} \leq \frac{rN_{k,\lfloor\frac{k}{2}\rfloor}}{N_{k-1,\lfloor\frac{k}{2}\rfloor-1}}$$

The RHS is bounded above for finite r. Since the ratio is also bounded above, its limit L always exists. Taking the limit as $k \to \infty$, we get

$$L = 2(1+r) - \frac{(r-1)^2}{L}.$$

This quadratic equation has two roots

$$L = 1 + r \pm 2\sqrt{r}.$$

To decide which sign we should pick, let $r = 1$, if we pick the minus sign $L = 0$ which is wrong given that the ratio is increasing and begins from $\frac{N_1(1)}{N_0(1)} = 1$.

2 Generalized matrix gamma distribution

Suppose there are N i.i.d. observations $x_1, x_2, \ldots, x_N$, coming from a p-dimensional multivariate normal distribution with an unknown covariance matrix. The likelihood has the following form:

$$p(\mathcal{D}|\mu, \Sigma) \propto |\Sigma|^{-\frac{N}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}(S_\mu \Sigma^{-1})\right),$$

where S_μ is the scatter matrix. Now, consider the following distribution, which we call "the generalized matrix gamma distribution":

$$GMG(\Sigma|S_0, n_0) \propto |\Sigma|^{-\frac{n_0-p-1}{2}} \exp\left(-\frac{1}{2\alpha}\operatorname{tr}[(S_0^{-1}\Sigma)^\alpha]\right), \tag{2.1}$$

where S_0 is a SPD matrix. We use it as the prior for the covariance matrix Σ. One can see that the posterior is

$$p(\Sigma|\mathcal{D}, \mu) \propto |\Sigma|^{-\frac{N+p+1-n_0}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}[S_\mu \Sigma^{-1} + \frac{1}{\alpha}(S_0^{-1}\Sigma)^\alpha]\right).$$

If we choose $n_0 = N + p + 1$ and set the parameter α to $\alpha = 3 - (1/t)$, then take the derivative of the logarithm of the above posterior kernel and equate it to zero, we have

$$-\Sigma^{-1}S_\mu\Sigma^{-1} + (S_0^{-1}\Sigma)^{2-\frac{1}{t}}S_0^{-1} = 0 \quad \longrightarrow \quad \Sigma = S_0(S_0^{-1}S_\mu)^t =: S_\mu \sharp_t S_0. \tag{2.2}$$

Therefore, the MAP estimation will be equal to the geodesic between the scatter matrix, S_μ, and the scale matrix, S_0. Note that we use the following proposition to compute the derivative:

Proposition 2.1. *Let A and X be two PSD matrices and $\alpha \in \mathbb{R}$. We have*

$$\frac{\partial \operatorname{tr}[(AX)^\alpha]}{\partial X} = \alpha(AX)^{\alpha-1}A. \tag{2.3}$$

2.1 Joint density of the eigenvalues

Let the PD matrix $\Sigma_{m\times m}$ has the following distribution

$$d\mathbb{P}(\Sigma) = \frac{1}{C(n,m,\alpha)}|\Sigma|^{\frac{n-m-1}{2}} \exp\Big(-\frac{1}{2\alpha}\operatorname{tr}(\Sigma^{\alpha})\Big)(d\Sigma), \tag{2.4}$$

where $C(n,m,\alpha)$ is the normalization constant. Employing the eigenvalue decomposition $\Sigma = Q\Lambda Q^T$, we have

$$d\mathbb{P}(\Lambda, Q) = \frac{1}{C(n,m,\alpha)}|Q\Lambda Q^T|^{\frac{n-m-1}{2}} \exp\Big(-\frac{1}{2\alpha}\operatorname{tr}[Q\Lambda^{\alpha}Q^T]\Big) \prod_{1\le i<j\le m}|\lambda_i-\lambda_j|\,(d\Lambda)(Q^T dQ).$$

Since $|Q\Lambda Q^T| = \prod_{i=1}^m \lambda_i$, and $\operatorname{tr}[Q\Lambda^{\alpha}Q^T] = \sum_{i=1}^m \lambda_i^{\alpha}$, the density can be written as follows:

$$d\mathbb{P}(\Lambda, Q) = \frac{1}{C(n,m,\alpha)}\prod_{i=1}^m \lambda_i^{\frac{n-m-1}{2}} \exp\Big(-\frac{1}{2\alpha}\lambda_i^{\alpha}\Big) \prod_{1\le i<j\le m}|\lambda_i-\lambda_j|\,(d\Lambda)(Q^T dQ).$$

Hence, the joint eigenvalue density is derived by integrating over $O(n)$, and taking care of overcounting:

$$d\mathbb{P}(\Lambda) = \frac{V_{\text{Orth}}(m)}{V_{\text{Phase}}(m)C(n,m,\alpha)}\prod_{i=1}^m \lambda_i^{\frac{n-m-1}{2}} \exp\Big(-\frac{1}{2\alpha}\lambda_i^{\alpha}\Big) \prod_{1\le i<j\le m}|\lambda_i-\lambda_j|\,(d\Lambda). \tag{2.5}$$

Therefore the normalization constant would be

$$C(n,m,\alpha) = \frac{2^{-m/2}\pi^{m(m+1)/4}}{\prod_{j=1}^m \Gamma(j/2)}\int_0^{\infty}\cdots\int_0^{\infty}\prod_{i=1}^m \lambda_i^{\frac{n-m-1}{2}} \exp\Big(-\frac{1}{2\alpha}\lambda_i^{\alpha}\Big) \prod_{1\le i<j\le m}|\lambda_i-\lambda_j|\,d\lambda_1\ldots\lambda_m. \tag{2.6}$$

For the general case of the generalized matrix gamma distribution, we must transform $\Sigma \mapsto S_0^{-1/2}\Sigma S_0^{-1/2}$ in (2.4):

$$GMG(\Sigma; S_0, n, \alpha) = \frac{1}{C(n,m,\alpha)}|S_0|^{-m}|S_0^{-1/2}\Sigma S_0^{-1/2}|^{\frac{n-m-1}{2}} \exp\Big(-\frac{1}{2\alpha}\operatorname{tr}[(S_0^{-1/2}\Sigma S_0^{-1/2})^{\alpha}]\Big). \tag{2.7}$$

Finally, because $\operatorname{tr}\big[(S_0^{-1}\Sigma)^{\alpha}\big] = \operatorname{tr}\big[(S_0^{-1/2}\Sigma S_0^{-1/2})^{\alpha}\big]$ and $\det(S_0^{-1}\Sigma) = \det(S_0^{-1/2}\Sigma S_0^{-1/2})$, we have

$$GMG(\Sigma; S_0, n, \alpha) = \frac{1}{C(n,m,\alpha)}|S_0|^{\frac{-n-m+1}{2}}|\Sigma|^{\frac{n-m-1}{2}} \exp\Big(-\frac{1}{2\alpha}\operatorname{tr}[(S_0^{-1}\Sigma)^{\alpha}]\Big). \tag{2.8}$$

2.2 Normalization constant

Proposition 2.2 (Selberg integral)**.**

$$\begin{aligned} S_N(\alpha,\beta,\gamma) &:= \int_0^1\cdots\int_0^1 \prod_{l=1}^N t_l^{\alpha}(1-t_l)^{\beta}\prod_{1\le j<k\le N}|t_k-t_j|^{2\gamma}\,dt_1\ldots dt_N \\ &= \prod_{j=0}^{N-1}\frac{\Gamma(\alpha+1+j\gamma)\,\Gamma(\beta+1+j\gamma)\,\Gamma(1+(j+1)\lambda)}{\Gamma(\alpha+\beta+2+(N+j-1)\gamma)\,\Gamma(1+\gamma)}. \end{aligned} \tag{2.9}$$

For calculating the integral part of $C(n,\alpha)$, one can use the procedure of Section 4.7.1 of [Forrester, 2010], which is applied to the Wishart distribution. For doing so, it is enough to have a closed-form expression for the following integral:

$$BS_N(\alpha,\beta,\kappa) := \int_0^1\cdots\int_0^1 \prod_{l=1}^N t_l^{\alpha}(1-t_l^{\kappa})^{\beta}\prod_{1\le j<k\le N}|t_k-t_j|\,dt_1\ldots dt_N. \tag{2.10}$$

We have $BS_N(\alpha, \beta, 1) = S_N(\alpha, \beta, 1/2)$, so this integral may have a Selberg-like closed form expression.

In the integral (2.10), if we change $t_l \mapsto \frac{t_l}{L}$, and put $N = m$, $\alpha = (n-m-1)/2$, $\beta = L^\kappa/\kappa$, then using the elementary limit $(1-(t_l/L)^\kappa)^{L^\kappa/\kappa} \to \exp(-t_l^\kappa/\kappa)$ as $L \to \infty$ gives

$$\frac{\Gamma_m(m/2)}{\pi^{m^2/2}} C(n, m, \kappa) = \lim_{L \to \infty} L^{\frac{m^2+n-1}{2}} \; BS_m\left(\frac{n-m-1}{2}, \frac{L^\kappa}{\kappa}, \kappa\right). \tag{2.11}$$

Question 2.3. *How to calculate the following integral for $\kappa > 0$:*

$$BS_N(\alpha, \beta, \kappa) := \int_0^1 \cdots \int_0^1 \prod_{l=1}^{N} t_l^\alpha (1 - t_l^\kappa)^\beta \prod_{1 \le j < k \le N} |t_k - t_j| \, dt_1 \ldots dt_N. \tag{2.12}$$

References

[Forrester, 2010] Forrester, P. J. (2010). *Log-gases and random matrices (LMS-34)*. Princeton University Press.

[Petersen, 2015] Petersen, T. K. (2015). Eulerian numbers. In *Eulerian Numbers*, pages 3–18. Springer.

[Sulanke, 2002] Sulanke, R. A. (2002). The narayana distribution. *Journal of statistical planning and inference*, 101(1-2):311–326.