Three Points Bounds for Determinants and Measures

An extension of the linear programming method for proving optimality of certain configurations, is the so-called 'three-point bounds'. In what follows $|\mathcal{C}| = N$ will be fixed and some potential function of three variables, $F(u, v, t) : [-1, 1)^3 \to \mathbb{R}$, will be considered, where $u = \langle x, y\rangle$, $v = \langle y, z\rangle$, and $t = \langle z, x\rangle$, and $x, y, z \in \mathcal{C} \subset \mathbb{S}^{d-1}$. It will be useful to adopt notation for the size of pre-image of a set of fixed triples (u, v, t) from the above correspondence with elements in $\mathcal{C}^3$ for $|\mathcal{C}| > 2$,

$$A_{u,v,t} = |\{(x, y, z) \in \mathcal{C}^3 \mid \langle x, y\rangle = u,\ \langle y, z\rangle = v,\ \text{and } \langle z, x\rangle = t\}|.$$

The primary question that is asked here is for lower bounds for the minimal energy of a configuration of size N in $\mathbb{S}^{d-1}$, where the energy is defined as

$$E_F(\mathcal{C}) = \sum_{\substack{x,y,z\in\mathcal{C}\\ x\neq y, y\neq z,\ z\neq x}} E(\langle x, y\rangle, \langle y, z\rangle, \langle z, x\rangle).$$

For an example of such an energy, one can rewrite the determinant-squared of a matrix spanned by vectors x, y, z on the sphere as

$$F(u, v, t) = \det\begin{pmatrix} 1 & u & v\\ u & 1 & t\\ v & t & 1\end{pmatrix} = 1 - u^2 - v^2 - t^2 + 2uvt.$$

Although it is an elementary problem to demonstrate that the energy corresponding to F is maximized by an ONB among size $N = 3$ collections of points on the sphere, we use the above potential as an illustration of the three-point bound method.

For three point bounds on the sphere in contrast to LP bounds infinite matrices replace the role of the orthogonal polynomials (Gegenbauer polynomials) which encode positive definiteness. These matrices, denoted M_k^d, have entries

$$M_k^d(i, j) = u^i v^j ((1 - u^2)(1 - v^2))^{\frac{k}{2}} P_k^{d-1}\left(\frac{t - uv}{\sqrt{(1 - u^2)(1 - v^2)}}\right).$$

The fact that these matrices are positive definite follows from the following result (taken from the form given in Cohn and Woo's 2011 paper).

Theorem 0.1. *Let H be a stabilizer in $O(d)$ of a point $e \in \mathbb{S}^{d-1}$. For each $k \geq 0$, there is an H-invariant matrix-valued positive-definite kernel on $\mathbb{S}^{d-1}$ taking $x_1, x_2 \in \mathbb{S}^{d-1}$ to the infinite matrix with (i_1, j_1) entry (starting at zero)*

$$u_1^{i_1} u_2^{j_1} ((1 - u_1)^2(1 - u_2)^2)^{\frac{k}{2}} P_k^{d-1}\left(\frac{t - u_1u_2}{\sqrt{(1 - u_1^2)(1 - u_2^2)}}\right)$$

where $u_j = \langle e, x_j\rangle$ an $t = \langle x_1, x_2\rangle$.

The role of positive-definiteness of such functions is analogous to that in LP bounds, as will be demonstrated in the following outline of the derivation of 3-point bounds given (with small adaptions) in Cohn-Woo for F as above.

By the above theorem, after symmetrizing the entries of M_k^d through averaging over permutations of variables u, v, and t, the resulting matrices, which will be called S_k^d, satisfy

$$\sum_{u,v,t} S_k^d(u, v, t) \succeq 0.$$

note that in the present case of $d = 3$, the entries of these matrices are generated from Chebyshev polynomials. It is now clear how analogous bounds to LP bounds might be developed here. For c constant, and F_k infinite positive semi-definite, set

$$H(u,v,t) = c + \sum_{k\geq 0} \langle F_k, S_k^d(u,v,t)\rangle.$$

Suppose now that H satisfies $H \leq F$ for $(u,v,t) \in D$, where

$$D = \{(u,v,t) \mid -1 \leq u,v,t < 1 \text{ and } 1 + 2uvt - u^2 - v^2 - t^2 \geq 0\}.$$

Then,

$$\sum_{(u,v,t)\in D} A_{u,v,t} H(u,v,t) \leq \sum_{(u,v,t)\in D} A_{u,v,t} F(u,v,t) = E_F(\mathcal{C}).$$

By positive semi-definiteness of $F_k,\ S_k^d$,

$$N(N-2)\langle F_k, J\rangle \delta_{k,0} + \sum_{(u,v,t)\in D} A_{u,v,t} \langle F_k, S_k^d(u,v,t)\rangle \geq 0$$

where J is the all-ones matrix, and so by summing over $k \geq 0$, one has

$$N(N-2)\langle F_0, J\rangle + \sum_{(u,v,t)\in D} A_{u,v,t}(H(u,v,t) - c) \geq 0.$$

Thus, combining these inequalities and noting that $\sum_{(u,v,t)\in D} A_{u,v,t} = N(N-1)(N-2)$,

$$E_F(\mathcal{C}) \geq N(N-2)[\langle F_0, J\rangle - (N-1)c]$$

holds for all $\mathcal{C} \subset \mathbb{S}^{d-1},\ |\mathcal{C}| = N$.

We now sketch an application of the above framework to the problem of maximizing E_F for $F(u,v,t) = \det \begin{pmatrix} 1 & u & v \\ u & 1 & t \\ v & t & 1 \end{pmatrix}$. For simplicity, we equivalently consider minimizing E_{-F} below.

Due to the expectations of the ONB attaining the minimum value of E_{-F}, it might suffice to restrict interest to only a few coefficients in the matrices F_k. Below, the entries will be restricted to the first three rows/columns in each matrix, F_0, F_1, and F_2, while simultaneously focusing only on entries (i,j) such that $i+j$ is even (since the potential function is symmetric on the sphere and is a low degree polynomial in u, v, and t). Additionally, F_0, F_1, and F_2 will be taken to be of the form $\begin{pmatrix} \alpha & 0 & \beta \\ 0 & 0 & 0 \\ \beta & 0 & \gamma \end{pmatrix}$, $\begin{pmatrix} \delta & 0 \\ 0 & \epsilon \end{pmatrix}$, and $\begin{pmatrix} 0 \end{pmatrix}$ respectively, where all entries beyond the last non-zero are not displayed (while it is understood that these matrices are infinite).

Many of the entries in F_k are zero, in aims of the lower bound for E_{-F} agreeing with the value on an orthonormal basis. Setting E_{-F} to zero (the value on an orthonormal basis) in the lower bound gives

$$3(1)[\langle F_0, J\rangle - (2)c] = 0 \quad \text{or equivalently} \quad c = \frac{\alpha + 2\beta + \gamma}{2}.$$

Under the above assumptions/considerations, the problem then of maximizing c subject to the conditions $H \leq F$, $(u,v,t) \in D$, $F_k \succeq 0$ may be rewritten as

$$
\max \frac{\alpha + 2\beta + \gamma}{2} \quad \text{s.t.} \quad
\begin{aligned}
\alpha + \frac{2}{3}\beta(u^2+v^2+t^2) + \frac{\gamma}{3}(u^2v^2+t^2u^2+v^2t^2) &\leq u^2+v^2+t^2-2uvt-1 \\
\alpha &\geq 0 \\
\gamma &\geq 0 \\
\alpha\gamma - \beta^2 &\geq 0 \\
-1 &\leq u,v,t < 1 \\
1+2uvt-u^2-v^2-t^2 &\geq 0
\end{aligned}
$$

.

The measure optimization problem for 3-point potentials may be treated similarly to the 2-point case. For an arbitrary probabilistic Borel measure μ we define its F-energy as

$$E_F(\mu) = \int\int\int F(\langle x,y\rangle, \langle y,z\rangle, \langle z,x\rangle) d\mu_x d\mu_y d\mu_z.$$

When the uniform measure over $\mathbb{S}^{d-1}$ is expected to be a minimizer of $E_F(\mu)$, it is more beneficial to use the semidefinite constraints as initially defined by Bachoc-Vallentin without simplifying the polynomials (see Theorem 3.2 in Bachoc-Vallentin):

$$Y_k^d(i,j) = P_i^{n+2k}(u) P_j^{n+2k}(v)((1-u^2)(1-v^2))^{\frac{k}{2}} P_k^{d-1}\left(\frac{t-uv}{\sqrt{(1-u^2)(1-v^2)}}\right).$$

Kernels $\langle A, Y_k^d\rangle$ are positive semidefinite for any $A \succeq 0$ and vanish over the uniform measure for any $k \geq 1$. The same is true for $k=0$ if we disregard the row 0 and column 0 of the Y_0^d matrix and use its submatrix for $i \geq 1$ and $j \geq 1$. When the optimization problem is formulated for measures over the projective space $\mathbb{RP}^{d-1}$, it is sufficient to consider the entries of Y_k^d satisfying the condition $i \equiv j \equiv k (\mod 2)$.

For the toy example of this approach, consider maximizing $E_F(\mu)$ for $F(u,v,t)$ defined above over projective measure on $\mathbb{RP}^2$, i.e.

$$F(u,v,t) = \det \begin{pmatrix} 1 & u & v \\ u & 1 & t \\ v & t & 1 \end{pmatrix} = 1 - u^2 - v^2 - t^2 + 2uvt.$$

Here are 2×2 submatrices of infinite matrices Y_0^3 (starting from row 1 and column 1), Y_1^3 and a 1×1 submatrix of Y_2^3, respectively:

$$\begin{pmatrix} uv & u\frac{3v^2-1}{2} \\ \frac{3u^2-1}{2}v & \frac{3u^2-1}{2}\frac{3v^2-1}{2} \end{pmatrix}, \begin{pmatrix} t-uv & u(t-uv) \\ v(t-uv) & uv(t-uv) \end{pmatrix}, \left(2(t-uv)^2 - (1-u^2)(1-v^2)\right).$$

In this scenario, it is sufficient to consider positive definite kernels generated by one element of Y_0^3, one element of Y_1^3 and one element of Y_2^3. Coefficients for the other elements of these submatrices should be 0 due to the observation about the case of projective spaces. Summing up the elements of Y_0^3, Y_1^3, and Y_2^3 over all six permutations of u, v, t, we get the following three symmetric positive definite kernels to use:

$$K_0 = \frac{1}{2}(9u^2v^2 + 9u^2t^2 + 9v^2t^2 - 6u^2 - 6v^2 - 6t^2 + 3);$$

$$K_1 = 6uvt - 2u^2v^2 - 2u^2t^2 - 2v^2t^2;$$

$$K_2 = -24uvt + 2u^2v^2 + 2u^2t^2 + 2v^2t^2 + 8u^2 + 8v^2 + 8t^2 - 6.$$

F can be represented as a linear combination of these kernels:

$$F(u,v,t) = 1 - u^2 - v^2 - t^2 + 2uvt = \frac{2}{9} - \frac{1}{27}K_0 - \frac{2}{9}K_1 - \frac{5}{36}K_2.$$

Integrating this equality we conclude

$$E_F(\mu) = \frac{2}{9} - \frac{1}{27}E_{K_0} - \frac{2}{9}E_{K_1} - \frac{5}{36}E_{K_2} \leq \frac{2}{9}.$$

It is fairly straightforward to check that the exact equality holds for the case of all isotropic measures.

Essentially the same proof works when maximizing E_F for measures in the d-dimensional space for $d > 3$. The submatrices to be used for semidefinite programming bounds are

$$\begin{pmatrix} uv & u\frac{dv^2-1}{d-1} \\ \frac{du^2-1}{d-1}v & \frac{du^2-1}{d-1}\frac{dv^2-1}{d-1} \end{pmatrix}, \begin{pmatrix} t-uv & u(t-uv) \\ v(t-uv) & uv(t-uv) \end{pmatrix}, \left(\frac{(d-1)(t-uv)^2-(1-u^2)(1-v^2)}{d-2}\right).$$

Similarly to the case $d = 3$, summing up the elements of Y_0^3, Y_1^3, and Y_2^3 over all six permutations of u, v, t, we get the following three symmetric positive definite kernels to use:

$$K_0 = \frac{2d^2}{(d-1)^2}(u^2v^2 + u^2t^2 + v^2t^2) - \frac{4d}{(d-1)^2}(u^2 + v^2 + t^2) + \frac{6}{(d-1)^2};$$

$$K_1 = 6uvt - 2(u^2v^2 + u^2t^2 + v^2t^2);$$

$$K_2 = -\frac{12(d-1)}{d-2}uvt + 2(u^2v^2 + u^2t^2 + v^2t^2) + \frac{2(d+1)}{d-2}(u^2 + v^2 + t^2) - \frac{6}{d-2}.$$

Again F can be represented as a linear combination of these kernels:

$$F(u,v,t) = 1-u^2-v^2-t^2+2uvt = \frac{(d-1)(d-2)}{d^2} - \frac{(d-1)(d-2)}{6d^2}K_0 - \frac{2(d-2)}{3d}K_1 - \frac{(3d-4)(d-2)}{6d(d-1)}K_2.$$

Integrating this equality we conclude

$$E_F(\mu) = \frac{(d-1)(d-2)}{d^2} - \frac{(d-1)(d-2)}{6d^2}E_{K_0} - \frac{2(d-2)}{3d}E_{K_1} - \frac{(3d-4)(d-2)}{6d(d-1)}E_{K_2} \leq \frac{(d-1)(d-2)}{d^2}.$$

Again it is fairly straightforward to check that E_{K_0}, E_{K_1}, E_{K_2} vanish simultaneously precisely for the case of all isotropic measures.

For a more general result, we consider the k-point potential $F_k(y_1, \dots, y_k)$ defined as the squared determinant of the Gram matrix of the set of vectors $\{y_1, \dots, y_k\}$. A measure μ does not have to be restricted to the surface of the unit sphere. Instead we impose a weaker normalizing condition: $\int \|x\|^2 d\mu_x = 1$. As in the previous result by a d-dimensional isotropic measure we mean μ satisfying $\int xx^* d\mu_x = \frac{1}{d} I_d$. It is easy to check that any isotropic measure satisfies the normalizing condition as well.

Similarly to previous questions, we define the F_k-energy for any normalized Borel probability measure μ in $\mathbb{R}^k$ as

$$E_k(\mu) = \int \dots \int F_k(x_1, \dots, x_k) d\mu_{x_1} \dots d\mu_{x_k}.$$

Theorem 0.2. *The set of maximizing measures of $E_k(\mu)$ in $\mathbb{R}^k$ is the set of isotropic measures in $\mathbb{R}^k$.*

Proof of D. Radchenko (slightly modified for a more general statement). It is sufficient to prove the statement for uniform distributions over a finite support. Assume the support of μ consists of N vectors $x_1, \dots, x_N$. Let D be a matrix formed by these vectors as columns. By $[N]_k$ we mean the set of all k-subsets of $\{1, \dots, N\}$ and by $[k]$ we mean $\{1, \dots, k\}$. For two sets of indices S_1 and S_2, D_{S_1,S_2} is a minor of D defined by the indices of S_1 for rows and the indices of S_2 for columns.

$$E_k(\mu) = \frac{k!}{N^k} \sum_{S \in [N]_k} det^2(D_{S,[k]}) = \frac{k!}{N^k} \sum_{S \in [N]_k} det(D_{S,[k]}) det(D^*_{[k],S})$$

By the Cauchy-Binet formula, this sum is just $det(DD^*)$, i.e.

$$E_k(\mu) = \frac{k!}{N^k} det(DD^*).$$

$tr(DD^*) = tr(D^*D) = N$ because of the normalizing condition. The sum of eigenvalues of DD^* is N and their product is not greater than $\left(\frac{N}{k}\right)^k$. Therefore,

$$E_k(\mu) = \frac{k!}{N^k} det(DD^*) \leq \frac{k!}{N^k} \left(\frac{N}{k}\right)^k = \frac{k!}{k^k}.$$

Equality is possible if and only if all eigenvalues of DD^* are $\frac{N}{k}$ which is precisely the condition for an isotropic measure. □