CPSC 536C: Algorithms for Convex Optimization
University of British Columbia

Convex Programs, Duality, John's Ellipsoid

Akshay Ram
January 25, 2025

These notes are subject to change and may contain errors. The proof of the Ellipsoid Lemma follows the notes of [Bubeck], as well as lectures of Daniel Dadush.

1 Introduction

The goal of this note is to prove the geometric update lemma for the ellipsoid method. This requires understanding the minimum containing ellipsoid for the set $\mathcal{E}_{t+1} \supseteq \mathcal{E}_t \cap H_t$, where $\mathcal{E}_t$ is the previous iterate ellipsoid, and H_t is the halfspace given by the separation oracle. We can in fact study this more generally: the problem of minimum containing ellipsoid for a convex set can be formulated as a convex program and analyzed using duality. So we first discuss the theory of convex duality; then we use it to prove John's Ellipsoid theorem, a fundamental result in convex geometry for minimum containing ellipsoids; finally we sketch a convex duality proof for the update lemma.

2 Convex Programming Duality

In this section, we consider the following framework for convex programs:

$$\inf_x f(x) \qquad \text{s.t.} \qquad \forall i \in [m] : f_i(x) \leq 0, \tag{1}$$

where f, f_i are all convex functions. It is often difficult to solve problems with constraints, so we would like to relax the problem by incorporating constraint functions into the objective.

Definition 1 (Lagrangian)**.** *For problem as in eq.* (1)*, let Lagrangian be*

$$\mathcal{L}(x, \lambda) := f(x) + \sum_{i \in [m]} \lambda_i f_i(x).$$

The dual function is $g(\lambda) := \inf_x \mathcal{L}(x, \lambda)$ and the dual program is

$$\sup_{\lambda \geq 0} g(\lambda) = \sup_{\lambda \geq 0} \left(\inf_x f(x) + \sum_{i \in [m]} \lambda_i f_i(x) \right). \tag{2}$$

Note that the dual problem is always a convex program even for non-convex f, f_i: g is concave as it is an inf of affine functions. The Lagrangian and dual give *lower bounds* on problem:

Fact 2 (Weak Duality)**.** *For any feasible x and dual $\lambda \geq 0$*

$$f(x) \geq L(x, \lambda) = f(x) + \sum_{i \in [m]} \lambda_i f_i(x) \geq \inf_x L(x, \lambda) = g(\lambda).$$

In general, there could be a *duality gap*, i.e. the value of the primal and dual programs could be different, so the lower bounds given by the dual program are not effective to understand the primal program. On the other hand, in certain cases there is no gap: for linear programs, strong duality always holds; for semi-definite programs, there are simple counterexamples with duality gap. In general applications, we will use the following sufficient condition:

Theorem 3 (Slater's Theorem)**.** *Consider the primal and dual programs eq.* (1)*, 2. If there is a strictly feasible solution x such that $\forall i : f_i(x) < 0$, then the primal and dual programs have equivalent values, and the supremum in the dual is attained.*

We will not prove this result here, but the result follows from the separating hyperplane theorem (as do basically all duality theorems). In fact we can even consider the more general convex program

$$\min_x f(x) \qquad \text{s.t.} \quad \forall i : f_i(x) \le 0, \quad Ax = b,$$

where we include linear equality constraints. In this setting, Slater's Theorem requires $f_i(x) < 0$ for all i and $Ax = b$, and still implies strong duality.

We next make some observations that can be deduced from optimality.

Proposition 4. *Let (x^*, λ^*) be primal/dual feasible solutions with $f(x^*) = g(\lambda^*)$. Then*

1. *x^* is feasible, i.e. $\forall i : f_i(x) \le 0$;*
2. *λ^* is feasible, i.e. $\lambda^* \ge 0$;*
3. *(Complementary Slackness): $\forall i : \lambda_i^* f_i(x^*) = 0$, i.e. λ^* is supported only on tight constraints;*
4. *(Lagrangian optimality): $x^* = \arg\inf_x L(x, \lambda^*)$.*

Proof: The first two properties are assumed. We deduce the remaining from the following sequence of inequalities:

$$g(\lambda^*) = f(x^*) \ge f(x^*) + \sum_i \lambda_i^* f_i(x^*) \ge g(\lambda^*),$$

where the first equality is by assumption, in the second we used primal/ dual feasibility so $\lambda_i^* f_i(x^*) \le 0$, and the final inequality was by definition of dual g. Therefore all inequalities must be tight: the first inequality being tight implies complementary slackness, and the last inequality being tight implies Lagrangian optimality. $\square$

This proposition allows us to reason about the optimizer, and applies even to non-convex programs. Of course in general there does not have to exist optimal solutions with matching primal and dual value, so we still need to check strong duality. But if somehow we have access to the optimizers of a general program, then the proposition allows us to deduce the above properties about the solution. We also note that (1), (2), (3) can be easily checked by evaluating the given functions, whereas (4) is more difficult in general. In the case of a convex (differentiable) program, we can use any of our optimality conditions for convex functions to verify Lagrangian optimality, which may be easier.

For convex programs, we also have the following converse.

Theorem 5 (KKT Conditions)**.** *Let eq.* (1)*, 2 be convex primal/dual programs, and assume (x^*, λ^*) satisfy (1), (2) primal/ dual feasibility; (3) complementary slackness; and (4) Lagrangian optimality, according to theorem 4. Then in fact (x^*, λ^*) are optimal solutions with $f(x^*) = g(\lambda^*)$.*

Proof: We follow the same sequence of inequalities, showing they are all tight:

$$f(x^*) = f(x^*) + \sum_i \lambda_i^* f_i(x^*) = g(\lambda^*),$$

the first step is complementary slackness $\forall i : \lambda_i^* f_i(x^*) = 0$, and the second is Lagrangian optimality. □

In this case we do not have to verify strong duality, as any primal and dual solutions satisfying the above conditions automatically certify strong duality. We also note that for differentiable convex programs, Lagrangian optimality is characterized by first-order condition

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) + \sum_i \lambda_i^* \nabla f_i(x^*) = 0,$$

which could be easy to directly verify if we can compute gradients.

We will use the first proposition to reason about the optimizer to show John's theorem; and we will use KKT conditions to certify optimality for the ellipsoid lemma.

3 John's Ellipsoid Theorem

The goal of this section is to prove the following:

Theorem 6. *Let $K \subseteq \mathbb{R}^n$ be a compact convex and full-dimensional set. Then there exists an ellipsoid $\mathcal{E}$ with center c such that*

$$c + \frac{1}{n}(\mathcal{E} - c) \subseteq K \subseteq \mathcal{E}.$$

Further, if K is origin-symmetric, i.e. $K = -K$, then this can be improved to

$$\frac{1}{\sqrt{n}}\mathcal{E} \subseteq K \subseteq \mathcal{E}.$$

Intuitively, we want to find the smallest ellipsoid containing K so that it is more likely for the scaled down version to fit inside K. For this we will use the volume as an objective function, which we will show is convex in the appropriate variables. We will also show that the containment condition can be written as convex constraints.

3.1 Convex Program Formulation

In the following, we let ellipsoid $\mathcal{E}$ be defined as

$$\mathcal{E} := c + AB_2^n$$

where $A \in \mathbb{R}^{n\times n}$ is invertible and $B_2^n := \{x \mid \|x\|_2 \leq 1\}$ is the standard Euclidean ball. We first rewrite containment in terms of our above representation. We leave the proof as an exercise.

Proposition 7. *Let $K := \mathrm{conv}\{x_i\}$; then $K \subseteq \mathcal{E}$ iff*

$$\forall i : \quad \langle (x_i - c), (AA^T)^{-1}(x_i - c)\rangle \leq 1.$$

Dually, consider inequality representation $K := \{x \in \mathbb{R}^n \mid Bx \leq d\}$; then $\mathcal{E} \subseteq K$ iff

$$\forall i : \quad \|A^T b_i\|_2 \leq d_i - \langle b_i, c\rangle.$$

We will eventually perform a change of variable so that these become convex constraints. Note that we have swapped containment between convex hull and inequality representation of K. This is for a very good reason: given $K := \operatorname{conv}\{x_i\}$, finding the maximum inscribed ellipsoid $\mathcal{E} \subseteq K$ is actually NP-hard! And similarly given $K := \{x \in \mathbb{R}^n \mid Bx \leq d\}$, finding minimum containing ellipsoid $K \subseteq \mathcal{E}$ is NP-hard.

Next we rewrite the objective function:

Fact 8. *For $\mathcal{E} = c + AB_2^n$ as above,*

$$vol(E) \propto |\det(A)| = \sqrt{\det(AA^T)}.$$

Proof: [Sketch] $\mathcal{E}$ is an affine transformation of B_2^n, so the constant in the proportion is $\operatorname{vol}(B_2^n)$. The first equality can be proven by integrating and noting the Jacobian is exactly $\det(A)$, and the second is by standard linear algebra. $\square$

Putting these together, we can construct a convex program for minimum containing ellipsoid using a change of variable:

Proposition 9. *Let $K := \operatorname{conv}\{x_i\}$; then the minimum containing ellipsoid is given by the optimizer of the following convex program:*

$$\min_{Q \succ 0} -\log\det(Q) \qquad s.t. \quad \forall i: \quad \langle x_i, Qx_i\rangle - 2\langle q, x_i\rangle + \langle q, Q^{-1}q\rangle \leq 1.$$

Proof: We first consider quadratic polynomial $f(x) := \langle x, Qx\rangle - 2\langle q, x\rangle + q_0$, where our variables are (Q, q, q_0). We want this to represent our ellipsoid $\mathcal{E} = \{x \mid f(x) \leq 1\}$ using change of variable $Q := (AA^T)^{-1}$. But if q_0 is unconstrained then we can always take it small enough to make any $Q \succeq 0$ feasible. Therefore we constrain it so that the quadratic is 0 at its center. Formally, we constrain $\min_x f(x) = 0$, which is equivalent to $q_0 = \langle q, Q^{-1}q\rangle$ as can be verified directly, noting $Q^{-1}q = \arg\min_x f(x)$. Finally, we claim that $(q, Q) \to \langle q, Q^{-1}q\rangle$ is jointly convex, and sketch the proof below: it is equivalent to show the epigraph is convex:

$$K := \left\{(q, Q, t) \mid \langle q, Q^{-1}q\rangle \leq t \iff \begin{pmatrix} Q & q \\ q^T & t \end{pmatrix} \succeq 0\right\},$$

where this last equivalence follows by Schur complement. Therefore the constraints are convex functions of (q, Q).

The objective comes from the volume calculation shown above, and it can be shown that $\log\det$ is a concave function for positive-definite input. $\square$

In the remainder, we show how optimality theorem 4 allow us to deduce John's theorem.

3.2 Symmetric Case

In this section we prove $\frac{1}{\sqrt{n}}\mathcal{E} \subseteq K \subseteq \mathcal{E}$ for symmetric $K = -K$. We focus on the minimum containing ellipsoid problem because it matches more directly with our proof of the ellipsoid lemma in the next section; but the maximum inscribed ellipsoid has the same (dual) proof. Note that we can assume $q = 0$ by symmetry and convexity: if (q, Q) is feasible then so is $(-q, Q)$ since $K = -K$, so we can assume $q = 0$ without loss of the objective value.

In order to use the theory of convex programming duality, we first need to verify strong duality. We use Slater's condition 3 with, for example $Q = R^{-2}I$ for any $R > \max_i \|x_i\|_2$. This gives a strictly feasible positive definite solution so strong duality and dual attainment holds.

Proposition 10. *Let $K \subseteq \mathbb{R}^n$ be compact, compact, full-dimensional, and symmetric $K = -K$. Then the minimum containing ellipsoid satisfies $\frac{1}{\sqrt{n}}\mathcal{E} \subseteq K \subseteq \mathcal{E}$.*

Proof: We write the KKT conditions for optimizers Q, λ: (1) and (2) are simple feasibility conditions; (3) complementary slackness tells us

$$\forall i : \lambda_i(\langle x_i, Qx_i\rangle - 1) = 0 \quad \text{i.e.} \quad (\lambda_i > 0 \implies \langle x_i, Qx_i\rangle = 1).$$

Geometrically, this implies that λ is supported on points x_i that lie on the boundary of the optimizing ellipsoid $\partial\mathcal{E} = \{x \mid \langle x, Qx\rangle = 1\}$. Finally (4) Lagrangian optimality gives

$$0 = \nabla_Q \mathcal{L}(Q, \lambda) = -Q^{-1} + \sum_i \lambda_i x_i x_i^T \implies Q^{-1} = \sum_i \lambda_i x_i x_i^T,$$

where we used $\nabla \log\det(Q) = Q^{-1}$, as can be verified by the cofactor expansion for determinant. Note that we implicitly assumed Q is invertible, i.e. $Q \succ 0$. But this follows from the non-degeneracy condition that K is full dimensional.

Using the above conditions, we want to argue that $\frac{1}{\sqrt{n}}\mathcal{E} \subseteq K$. We will argue the contrapositive, so given $y \notin K$ we want to show $y \notin \frac{1}{\sqrt{n}}\mathcal{E}$, i.e. $\langle y, Qy\rangle > 1/n$, since $r\mathcal{E} = \{x \mid \langle x, Qx\rangle \leq r^2\}$.

In this section we will mostly perform algebraic manipulations to show the result. In the following non-symmetric case, we will consider a slightly more geometric proof.

So assuming $y \notin K$, we must have a separating hyperplane

$$\exists w : \quad \langle w, y\rangle > \sup_{x \in K} \langle w, x\rangle = \max_i |\langle w, x_i\rangle,$$

where we used that $K = \text{conv}\{x_i\}$ and $K = -K$ is symmetric so we can take the absolute value. Continuing, we have

$$\langle y, Qy\rangle = \|Q^{1/2}y\|_2^2 \geq \frac{\langle Q^{-1/2}w, Q^{1/2}y\rangle^2}{\|Q^{-1/2}w\|_2^2} = \frac{\langle w, y\rangle^2}{\langle w, Q^{-1}w\rangle},$$

where we used that $\|z\|_2 = \max_{\|u\|_2 \leq 1}\langle u, z\rangle$. Now we upper bound the denominator

$$\langle w, Q^{-1}w\rangle = \sum_i \lambda_i \langle w, x_i\rangle^2 < \sum_i \lambda_i \langle w, y\rangle^2,$$

where the first step was by optimality $Q^{-1} = \sum_i \lambda_i x_i x_i^T$, and in the final step we used that w is a separating hyperplane. Finally we bound the dual

$$\sum_i \lambda_i \cdot 1 = \sum_i \lambda_i \langle x_i, Qx_i\rangle = \langle Q^{-1}, Q\rangle = n,$$

where the first step was by complementary slackness, $\lambda_i > 0 \implies \langle x_i, Qx_i\rangle = 1$, the second was by optimality for Q again.

Putting this altogether gives our result:

$$y \notin K \implies \langle y, Qy\rangle > \frac{\langle w, y\rangle^2}{\langle w, y\rangle^2}\left(\sum_i \lambda_i\right)^{-1} = \frac{1}{n} \implies y \notin \frac{1}{\sqrt{n}}\mathcal{E}.$$

3.3 General Case

In this section we prove the general case of John's theorem. The proof plan is much the same, but a few of the steps are slightly more complicated. We can verify Slater's condition the same way, by choosing a large enough $R > \max_i \|x\|_2$ so that $(0, R^{-2}I)$ is strictly feasible, and we have strong duality by theorem 3. In the remainder we consider optimality conditions and prove the following:

Proposition 11. *Let $K \subseteq \mathbb{R}^n$ be compact, convex, and full-dimensional. Then the minimum containing ellipsoid $\mathcal{E}$ with center c satisfies*

$$c + \frac{1}{n}(\mathcal{E} - c) \subseteq K \subseteq \mathcal{E}.$$

Proof: We consider KKT conditions for optimizers (q, Q, λ): primal feasibility is given in theorem 9, dual feasibility is simply $\lambda \geq 0$; for complementary slackness we have

$$\lambda_i > 0 \implies \quad \langle (x_i - c), Q(x_i - c)\rangle = 0,$$

where $c := Q^{-1}q$ is the center, i.e. x_i is at the boundary of the ellipsoid; and finally Lagrangian optimality gives:

$$\nabla_q \mathcal{L}(q, Q, \lambda) = 2\sum_i \lambda_i(Q^{-1}q - x_i) = 0 \iff \sum_i \lambda_i x_i = \sum_i \lambda_i c$$

$$\nabla_Q \mathcal{L}(q, Q, \lambda) = -Q^{-1} + \sum_i \lambda_i(x_i x_i^T - Q^{-1}qq^TQ^{-1}) = 0 \iff Q^{-1} = \sum_i \lambda_i (x_i - c)(x_i - c)^T,$$

where in both calculations we use $c := Q^{-1}q$ for the center, and in the second we use that $\sum_i \lambda_i x_i = \sum_i \lambda_i c$ from q-optimality, so we can replace $\sum_i \lambda_i(x_i x_i^T - cc^T) = \sum_i (x_i - c)(x_i - c)^T$. Next we perform a transformation to simplify notation. Note that all containment relations are invariant under the affine transformation $x \to z := Q^{1/2}(x - c)$. This transformation shifts and scaled our body to $K \to K' := \mathrm{conv}\{z_i\}$, and the minimum containing ellipsoid becomes the Euclidean ball $K' \subseteq B_2^n$. Further, the optimality conditions become

$$c = 0, \quad \sum_i \lambda_i z_i = \sum_i \lambda_i c = 0, \quad I = \sum_i \lambda_i z_i z_i^T,$$

and complementary slackness is $\lambda_i > 0 \implies \|z_i\|_2^2 = 1$.

The remainder follows the original proof of Fritz John (Extremum problems with inequalities as subsidiary conditions). The goal is now to show that $\frac{1}{n}B_2^n \subseteq K'$. For this we will show that $h_{K'}(u) = \sup_{z \in K'} \langle u, z\rangle \geq \|u\|_2/n$ for all u; this implies our containment relation by polar duality:

$$\frac{1}{n}B_2^n \subseteq K' \iff (K')^\circ \subseteq (\frac{1}{n}B_2^n)^\circ \iff \forall u : \sup_{z \in K'} \langle u, z\rangle \geq \frac{1}{n} \sup_{z \in B_2^n} \langle u, z\rangle = \frac{\|u\|_2}{n}.$$

We require two further observations: first note that $K' \subseteq B_2^n$ implies $h_{K'}(u) \leq \|u\|_2$; and second, we can prove $\sum_i \lambda_i = n$ by complementary slackness, just as in the symmetric case.

Now let $u \in S^{n-1}$ and $h_{K'}(u) =: \mu$. Then we must have $\forall i : \langle u, z_i\rangle \in [-1, \mu]$ iff $(\langle u, z_i\rangle + 1)(\langle u, z_i\rangle - \mu) \leq 0$. Now we can use the dual certificate to show the required bound:

$$0 \geq \sum_i \lambda_i(\langle u, z_i\rangle + 1)(\langle u, z_i\rangle - \mu) = \sum_i \lambda_i \langle u, z_i\rangle^2 + (1-\mu)\sum_i \lambda_i \langle u, z_i\rangle - \sum_i \lambda_i \mu = \|u\|_2^2 + 0 - n\mu,$$

where we used optimality conditions $\sum_i \lambda_i z_i z_i^T = I, \sum_i \lambda_i z_i = 0$ and $\sum_i \lambda_i = n$. Rearranging, this gives $\mu \geq 1/n$, and since $u \in S^{n-1}$ was arbitrary, this shows $\forall u : h_{K'}(u) \geq \|u\|_2/n$. $\square$

4 Ellipsoid Lemma

We restate the required result for the Ellipsoid algorithm:

Lemma 12 (Lemma 2.3 in Bubeck)**.** *Let $\mathcal{E} := \{x \in \mathbb{R}^n \mid \langle (x - c_p), P(x - c_p)\rangle \leq 1\}$ be an Ellipsoid with $P \succ 0$; and let $H := \{x \in \mathbb{R}^n \mid \langle w, x\rangle \leq \langle w, c_p\rangle\}$ be a halfspace through the center of $\mathcal{E}$. Then there exists ellipsoid $\overline{\mathcal{E}}$ such that*

$$\overline{\mathcal{E}} \supset \mathcal{E} \cap H \quad \textit{and} \quad \mathit{vol}(\overline{\mathcal{E}}) \leq \mathit{vol}(\mathcal{E}) \exp(-1/2n).$$

Further, if $n \geq 2$ then $\overline{\mathcal{E}} := \{x \in \mathbb{R}^n \mid \langle (x - c), Q(x - c)\rangle \leq 1\}$ where

$$c := c_p - \frac{1}{n+1} \frac{P^{-1}w}{\sqrt{\langle w, P^{-1}w\rangle}}, \qquad Q := \left(1 - \frac{1}{n^2}\right) P + \frac{2(n+1)}{n^2} \frac{ww^T}{\langle w, P^{-1}w\rangle}.$$

We first verify that the proposed ellipsoid satisfies the volume bound. Then we show how to derive this update using minimum containing ellipsoid and convex duality. For the remainder, we perform an affine transformation to simplify notation: for $x \to P^{1/2}(x - c_p)$, the original ellipsoid becomes $\mathcal{E} = B_2^n$ with center $c_p = 0$, and without loss we can assume the halfspace is defined by normal vector $\|w\|_2 = 1$.

Proposition 13. *For large enough n update ellipsoid given in theorem 12 satisfies*

$$\mathit{vol}(\overline{\mathcal{E}}) \leq \mathit{vol}(\mathcal{E}) \exp(-\Omega(1/n)).$$

Proof: We rewrite the update after the affine transformation:

$$c = -\frac{1}{n+1} w, \qquad Q = \left(1 - \frac{1}{n^2}\right) I + \frac{2(n+1)}{n^2} ww^T.$$

Recall the volume of the ellipsoid is proportional to the inverse square root of the determinant, so we want to lower bound

$$\det(Q) = \left(1 - \frac{1}{n^2}\right)^{n-1} \left(1 - \frac{1}{n^2} + \frac{2(n+1)}{n^2}\right) \approx \exp\left(\frac{-(n-1) + 2n + 1}{n^2}\right) \geq \exp\left(\frac{1}{n}\right),$$

where we used the heuristic Taylor approximation $1 + x \approx e^x$ for small x. The exact calculation is performed properly in the proof in [Bubeck]. $\square$

In the remainder, we sketch how the KKT conditions for minimum containing ellipsoid allow us to derive and certify the update. We follow the notes of Daniel Dadush. In particular, recall the convex program

$$\min -\log\det(Q) \qquad \text{s.t.} \quad \forall x \in \mathcal{E} \cap H : \quad q(x) := \langle (x - c), Q(x - c)\rangle \leq 1. \tag{3}$$

We first observe that our problem is symmetric around the w axis since $\mathcal{E} = B_2^n$. Therefore if q is feasible, then so is any rotation around w, and so by convexity we can without loss assume the solution is symmetric around the w axis. Next, we observe that $p(x) := \langle x, Ix\rangle$, i.e. $c = 0, Q = I$, is a feasible solution as $B_2^n \supseteq B_2^n \cap H$. We also consider another feasible solution given

$$\delta(x) := \langle (x - \frac{1}{2}w), (4ww^T)(x - \frac{1}{2}w)\rangle = 4(\langle x, w\rangle - \frac{1}{2})^2.$$

This defines the degenerate ellipsoid $\{x \mid \langle x, w\rangle \in [0,1]\}$ and clearly also contains $B_2^n \cap H$. By convexity, this implies $q_\gamma := (1-\gamma)p + \gamma\delta$ is also a feasible solution for all $\gamma \in [0,1]$. We can calculate the center and matrix

$$c_\gamma = (1-\gamma)\cdot 0 + \frac{\gamma}{2}w = \frac{\gamma}{2}w, \quad Q_\gamma = (1-\gamma)I + 4\gamma ww^T.$$

This matches the symmetry structure that we expect from our solution.

Finally, we write down the KKT conditions where we used $K := B_2^n \cap H$ for shorthand. Note that we have infinite dual variables $\{\lambda_x\}_{x\in K}$ (we sweep these technicalities under the rug), so we use integrals instead of sums:

$$\nabla = 0 \implies \quad \int_{x\in K} \lambda_x (x-c) = 0; \quad Q^{-1} = \int_{x\in K} \lambda_x (x-c)(x-c)^T;$$

$$(CS): \quad \operatorname{supp}(\lambda) \subseteq \{x \mid q(x) = 1\}.$$

For the candidate primal solution q_γ, if we can verify feasibility and find a dual solution satisfying the above, then theorem 5 proves optimality. We can interpret the KKT conditions as looking for a measure λ on contact points with mean c and covariance Q^{-1}. We next show how to construct a family of measures meeting these criteria for (c_γ, Q_γ): for each α, let λ^α put α probability at w and the remaining $1-\alpha$ probability uniformly on $S^{n-1} \cap \partial H$, i.e. the great circle at the bound of the intersection $B_2^n \cap H$. We can calculate mean and covariance:

$$\mathbb{E}_{\lambda^\alpha}[x] = \alpha w + (1-\alpha)\int_{S^{n-1}\cap\partial H} x = \alpha w =: c,$$

$$\mathbb{E}_{\lambda^\alpha}[xx^T] - cc^T = \alpha ww^T + (1-\alpha)\int_{S^{n-1}\cap\partial H} xx^T - \alpha^2 ww^T = \alpha(1-\alpha)ww^T + \frac{1-\alpha}{n-1}(I - ww^T).$$

This roughly matches the family $(c_\gamma, Q_\gamma^{-1})$. To finish, we need to check which value of γ gives primal feasibility and which value of α gives complementary slackness. We leave these verifications as exercises for the reader.