SELECTED HINTS, SECTION 2.3

Problems 1, 2 (after having drawn the picture), and 5 are pretty straightforward. Problem 7 isn't:

7. Consider a matrix A in $\mathbb{S}^n_{++}$ and a real $b > 0$.
   (a) Assuming the problem
   $$\inf\{-\log\det X : \operatorname{tr} AX \le b,\ X \in \mathbb{S}^n_{++}\}$$
   has a solution, find it.
   (b) Repeat using the objective function $\operatorname{tr} X^{-1}$.
   (c) Prove the problems in parts (a) and (b) have optimal solutions. (Hint: Section 1.2, Exercise 14.)

(a) In principle, this exercise is straightforward: Try to solve the KKT conditions. The "only" difficulty is the computation of the gradients...

Recall that the gradient of a differentiable function $h\colon \mathbb{S}^n_{++} \to \mathbb{R}$ at $X \in \mathbb{S}^n_{++}$ is defined to be the symmetric matrix $\nabla h(X)$ satisfying the equation
$$h'(X;Y) = \langle \nabla h(X), Y\rangle$$
for all $Y \in \mathbb{S}^n$.

- The computation of the gradient of the constraint function should be completely straightforward.
- Computation of $\nabla f(X)$ with $f(X) = -\log\det X$ (this is a computation every mathematician should have done at least [or: exactly...] once in his life):
  Denote $h(X) := \det X$. Finding $\nabla f(X)$ is then trivial once we know $\nabla h(X)$.
  For the computation of $\nabla h(X)$, try first to show that $\nabla h(\mathrm{Id}) = \mathrm{Id}$ ($\mathrm{Id} \in \mathbb{R}^{n\times n}$ denotes the identity matrix). This can, for instance, be done by computing directional derivatives of the determinant in direction of elementary matrices (matrices with only a single non-zero entry). Then it is somehow simple to deduce $\nabla h(X)$ for arbitrary X using the properties of determinants.

(b) Here one ends up with having to compute the gradient of the function $f(X) = \operatorname{tr} X^{-1}$. One possibility for approaching this problem is to compute directional derivatives of f. A valuable tool for this computation can be the Neumann series expansion
$$(\mathrm{Id} - Y)^{-1} = \sum_k Y^k,$$
valid whenever $\|Y\|$ is sufficiently small.

(c) Use the fact that a constrained optimization problem $\inf f(x)$ such that $g(x) \le 0$ with f, g continuous has a solution if all the level sets of $f(x)+g(x)$ are compact (try to prove this!).

For proving the compactness of the level sets of the respective functions $f + g$, one mainly has to find (or rather: prove the existence of) upper and lower bounds for the eigenvalues of matrices X with $f(X) + g(X) \le c$ for given $c \in \mathbb{R}$. Here it will be necessary to derive a *lower* bound for $\operatorname{tr}(AX)$ in

Date: October 14, 2015.

terms of $\lambda(A)$ and $\lambda(X)$. This can be achieved by applying Fan's inequality to A and $-X$.