Exercise 0.1 (Primal and dual problem in the sense of optimization theory). Suppose $C \in \mathrm{End}(\mathbf{R}^p)$ to be symmetric and positive definite; that is, $\langle\, Cy, y\,\rangle = \langle\, y, Cy\,\rangle$ and $\langle\, y, Cy\,\rangle \geq 0$ for all $y \in \mathbf{R}^p$, with equality only if $y = 0$. Furthermore, let $n \leq p$ and suppose $A \in \mathrm{Lin}(\mathbf{R}^n, \mathbf{R}^p)$ to be injective.

(i) Prove that $C \in \mathrm{Aut}(\mathbf{R}^p)$ and that $A^tCA \in \mathrm{End}(\mathbf{R}^n)$ is symmetric and positive definite, and therefore satisfies $A^tCA \in \mathrm{Aut}(\mathbf{R}^n)$. (Recall that $A^t \in \mathrm{Lin}(\mathbf{R}^p, \mathbf{R}^n)$ is defined by $\langle\, A^t y, x\,\rangle = \langle\, y, Ax\,\rangle$, for all $y \in \mathbf{R}^p$ and $x \in \mathbf{R}^n$.)

Let $0 \neq a \in \mathbf{R}^n$ be fixed and define the quadratic function

$$P : \mathbf{R}^n \to \mathbf{R} \qquad \text{by} \qquad P(x) = \frac{1}{2}\langle\, A^tCAx, x\,\rangle - \langle\, a, x\,\rangle.$$

(ii) For $x \in \mathbf{R}^n$, show by means of part (i) that $DP(x) = 0$ if and only if x satisfies the linear equation $A^tCAx = a$ and that such an x is unique. Conclude that P attains the value $p := -\frac{1}{2}\langle\, a, (A^tCA)^{-1}a\,\rangle$ at its only critical point.

In the sequel it may be used without proof that $\min_{x\in\mathbf{R}^n} P(x) = p$. (This fact can be proved using compactness and consideration of the asymptotic behavior of $P(x)$ as $\|x\| \to \infty$.)

Now we come to the main issue of the exercise, namely, the study of the quadratic function

$$Q : \mathbf{R}^p \to \mathbf{R} \qquad \text{given by} \qquad Q(y) = \frac{1}{2}\langle\, C^{-1}y, y\,\rangle, \qquad \text{under the constraint} \qquad A^t y = a.$$

(iii) Demonstrate that, for all $y \in V := \{\, y \in \mathbf{R}^p \mid A^t y = a\,\}$ and $x \in \mathbf{R}^n$, we have the following identity, in which an *uncoupled* expression occurs at the left-hand side,

$$Q(y) + P(x) = \frac{1}{2}\langle\, C(C^{-1}y - Ax),\, C^{-1}y - Ax\,\rangle.$$

Deduce, for $y \in V$ and $x \in \mathbf{R}^n$, that we have $Q(y) \geq -P(x)$, with equality if and only if $y = CAx$. Using part (ii), show, for all $y \in V$,

$$Q(y) \geq -p = \max_{x\in\mathbf{R}^n} -P(x), \qquad \text{and conclude} \qquad \min_{y\in V} Q(y) = \max_{x\in\mathbf{R}^n} -P(x).$$

In other words, the constrained minimum of Q equals the unconstrained maximum of $-P$. As an example of a different approach, we now study the preceding problem by introducing the Lagrange function

$$L : \mathbf{R}^p \times \mathbf{R}^n \to \mathbf{R} \qquad \text{with} \qquad L(y, x) = Q(y) - \langle x, (A^t y - a)\rangle.$$

(iv) Using L, determine the points $y \in V$ where the extrema of $Q|_V$ are attained and derive the same results as in part (iii).

Background. The result above is one of the simplest cases of a duality that plays an important role in *optimization theory*. In this manner, the *primal problem* of minimizing Q under constraints is replaced by the *dual problem* of maximizing P.

Solution of Exercise 0.1

(i) Suppose that $Cy = 0$, then $\langle\, y,\, Cy\,\rangle = 0$, hence $y = 0$. Accordingly, C is injective and thus $C \in \mathrm{Aut}(\mathbf{R}^p)$. Next, $(A^tCA)^t = A^tC^tA^{tt} = A^tCA$, which proves the symmetry. Further, assume $x \in \mathbf{R}^n$ satisfies $A^tCAx = 0$. Then, in view of C being positive definite and A injective,

$$\langle\, x,\, A^tCAx\,\rangle = \langle\, Ax,\, CAx\,\rangle = 0 \quad\Longrightarrow\quad Ax = 0 \quad\Longrightarrow\quad x = 0.$$

Finally, apply the first argument to A^tCA.

(ii) The first assertion on $DP(x)$ follows from Corollary 2.4.3.(ii), while the uniqueness of x is a consequence of $A^tCA \in \mathrm{Aut}(\mathbf{R}^n)$. Furthermore,

$$P((A^tCA)^{-1}x) = \frac{1}{2}\langle\, a, (A^tCA)^{-1}a\,\rangle - \langle\, a, (A^tCA)^{-1}a\,\rangle.$$

(iii) For all $y \in V$ and $x \in \mathbf{R}^n$ one obtains, using $A^ty = a$ and the positive definiteness of C,

$$\begin{aligned} &Q(y) + P(x)\\ &\quad= \frac{1}{2}\langle\, C^{-1}y, y\,\rangle + \frac{1}{2}\langle\, A^tCAx, x\,\rangle - \langle\, a, x\,\rangle\\ &\quad= \frac{1}{2}\langle\, C(C^{-1}y), C^{-1}y\,\rangle + \frac{1}{2}\langle\, CAx, Ax\,\rangle - \langle\, A^ty, x\,\rangle\\ &\quad= \frac{1}{2}\langle\, C(C^{-1}y - Ax), C^{-1}y - Ax\,\rangle + \frac{1}{2}\langle\, y, Ax\,\rangle + \frac{1}{2}\langle\, CAx, C^{-1}y\,\rangle - \langle\, y, Ax\,\rangle\\ &\quad= \frac{1}{2}\langle\, C(C^{-1}y - Ax),\, C^{-1}y - Ax\,\rangle \geq 0. \end{aligned}$$

Once more on the basis of C being positive definite, one has equality if and only if $C^{-1}y - Ax = 0$, in other words, $y = CAx$. In turn, this implies $Q(y) \geq -P(x)$, for all $y \in V$ and $x \in \mathbf{R}^n$. In particular, this is the case if $x^0 \in \mathbf{R}^n$ is the unique element satisfying $A^tCAx^0 = a$ (see part (ii)); this implies, for all $y \in V$,

$$Q(y) \geq -P(x^0) = \max_{x\in\mathbf{R}^n} -P(x) = -\min_{x\in\mathbf{R}^n} P(x) = -p.$$

Now consider $y^0 = CAx^0 \in \mathbf{R}^p$. Then $A^ty^0 = A^tCAx^0 = a$, that is, $y^0 \in V$; and the preceding arguments imply $Q(y^0) = -P(x^0) = -p$. This proves $\min_{y\in V} Q(y) = -p$.

(iv) Applying the method of Lagrange multipliers, one obtains that extrema for $Q|_V$ occur at points $y \in V$ satisfying

$$D_yL(y,x) = C^{-1}y - Ax = 0 \quad\Longrightarrow\quad y = CAx \qquad\text{and}\qquad a = A^ty = A^tCAx.$$

However, for such y and x,

$$\begin{aligned} Q(y) &= \frac{1}{2}\langle\, C^{-1}CAx, CAx\,\rangle = \frac{1}{2}\langle\, Ax, CAx\,\rangle = \frac{1}{2}\langle\, A^tCAx, x\,\rangle\\ &= -\frac{1}{2}\langle\, A^tCAx, x\,\rangle + \langle\, a, x\,\rangle = -P(x). \end{aligned}$$

C^{-1} being positive definite implies that Q attains a minimum on V; indeed, the graph of the restriction of Q to V is the intersection of an elliptic paraboloid and an affine submanifold (if necessary, use that continuity of the function Q implies that it attains extrema on compact subsets of V). Therefore $\min_{y\in V} Q(y) = -P(x)$ where $x = (A^tCA)^{-1}a \in \mathbf{R}^n$. Finally, use part (ii) to obtain the desired equality.

Background. The method of Lagrange multipliers enables one to obtain the dual quadratic form P, given the primal form Q together with its constraint, by explicitly computing the minimal value of Q.