Thus, we obtain a simple expression for $\|w\|^2$ in terms of λ and μ.

The vectors u_i and v_j for which the i-th inequality is active and the $(p+j)$th inequality is active are called *support vectors*. For every vector u_i or v_j that is not a support vector, the corresponding inequality is inactive, so $\lambda_i = 0$ and $\mu_j = 0$. Thus we see that *only the support vectors contribute to a solution.* If we can *guess* which vectors u_i and v_j are support vectors, namely, those for which $\lambda_i \neq 0$ and $\mu_j \neq 0$, then for each support vector u_i we have an equation

$$-\sum_{j=1}^{p} u_i^\top u_j \lambda_j + \sum_{k=1}^{q} u_i^\top v_k \mu_k + b + 1 = 0,$$

and for each support vector v_j we have an equation

$$\sum_{i=1}^{p} v_j^\top u_i \lambda_i - \sum_{k=1}^{q} v_j^\top v_k \mu_k - b + 1 = 0,$$

with $\lambda_i = 0$ and $\mu_j = 0$ for all non-support vectors, so together with the Equation $(*_2)$ we have a linear system with an equal number of equations and variables, which is solvable if our separation problem has a solution. Thus, in principle we can find λ, μ, and b by solving a linear system.

Remark: We can first solve for λ and μ (by eliminating b), and by $(*_1)$ and since $w \neq 0$, there is a least some nonzero λ_{i_0} and thus some nonzero μ_{j_0}, so the corresponding inequalities are equations

$$\begin{aligned}
-\sum_{j=1}^{p} u_{i_0}^\top u_j \lambda_j + \sum_{k=1}^{q} u_{i_0}^\top v_k \mu_k + b + 1 &= 0 \\
\sum_{i=1}^{p} v_{j_0}^\top u_i \lambda_i - \sum_{k=1}^{q} v_{j_0}^\top v_k \mu_k - b + 1 &= 0,
\end{aligned}$$

so b is given in terms of λ and μ by

$$b = \frac{1}{2}(u_{i_0}^\top + v_{j_0}^\top)\left(\sum_{i=1}^{p} \lambda_i u_i - \sum_{j=1}^{p} \mu_j v_j\right).$$

Using the dual of the Lagrangian, we can solve for λ and μ, but typically b is not determined, so we use the above method to find b.

The above nondeterministic procedure in which we guess which vectors are support vectors is not practical. We will see later that a practical method for solving for λ and μ consists in maximizing the dual of the Lagrangian.