

**Due: 17 Sept 2020**

1. **KKT matrix.** Show that the KKT matrix of the system

$$\begin{bmatrix} G & A_W^\top \\ A_W & 0 \end{bmatrix} \begin{bmatrix} -p_k \\ \lambda \end{bmatrix} = \begin{bmatrix} \nabla f(x_k) \\ 0 \end{bmatrix}.$$

with  $G \in \mathbb{R}^{d \times d}$  SPD,  $A \in \mathbb{R}^{m \times d}$  with linearly independent rows, is of *saddle-point type*. That is, it has  $d$  positive eigenvalues and  $m$  negative eigenvalues. *Hint: omit the subscript  $W$  for brevity. Find matrices  $X$  and  $S$  ( $S$  is called the **Schur complement**) such that*

$$\begin{bmatrix} G & A^\top \\ A & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} G & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & X^\top \\ 0 & I \end{bmatrix}.$$

Then use Sylvester's Law of Inertia to finish the proof.

**Solution.** We begin by multiplying out the matrix equation.

$$\begin{aligned} \begin{bmatrix} G & A^\top \\ A & 0 \end{bmatrix} &= \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} G & 0 \\ 0 & S \end{bmatrix} \begin{bmatrix} I & X^\top \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} G & 0 \\ XG & S \end{bmatrix} \begin{bmatrix} I & X^\top \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} G & GX^\top \\ XG & XGX^\top + S \end{bmatrix} \end{aligned}$$

Equating both sides gives the following four equations:

$$\begin{aligned} G &= G \\ GX^\top &= A^\top \\ XG &= A \\ XGX^\top + S &= 0 \end{aligned}$$

which imply that  $X = AG^{-1}$  ( $G$  is symmetric so it is invertible) and hence that

$$\begin{aligned} 0 &= XGX^\top + S = (AG^{-1})G(AG^{-1})^\top + S = AG^{-1}A^\top + S \\ \implies S &= -AG^{-1}A^\top \end{aligned}$$

where we also make use of  $G$  being symmetric to write  $G^{-\top} = G^{-1}$ . Rewriting our original matrix equation now gives

$$\begin{bmatrix} G & A^\top \\ A & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ AG^{-1} & I \end{bmatrix} \begin{bmatrix} G & 0 \\ 0 & -AG^{-1}A^\top \end{bmatrix} \begin{bmatrix} I & G^{-1}A^\top \\ 0 & I \end{bmatrix}.$$

Since  $G \in \mathbb{R}^{d \times d}$  is symmetric, it has  $d$  distinct eigenvalues and since  $G$  is positive-definite all the eigenvalues are positive. Since  $A \in \mathbb{R}^{m \times d}$  has full rank, it is a surjective map onto its image, i.e. its reduced row echelon form has all 1's down the main diagonal. Let  $G^{-1} = QDQ^\top$  with  $Q \in \mathbb{R}^{d \times d}$  an orthogonal matrix and  $D = \text{diag}\{1/\lambda_i\}$  where  $\lambda_i > 0$  are the eigenvalues of  $G$ ; we can do all of this as  $G$  is SPD. Then we have

$$\begin{aligned} AG^{-1} &= AQDQ^\top \\ AG^{-1}A^\top &= AQDQ^\top A^\top \\ -AG^{-1}A^\top &= AQ(-D)Q^\top A^\top. \end{aligned}$$

Since  $A$  has full rank and  $Q$  is orthogonal, we can rewrite them into  $B = AQ$  which will also have full rank. Hence  $-AG^{-1}A^\top = B(-D)B^\top \in \mathbb{R}^{m \times m}$ , which is where the  $m$  negative eigenvalues come from. (As  $A$  is full rank, none of the eigenvalues of  $G^{-1}$  are mapped to zero, implying that there are  $m$  negative eigenvalues.) To conclude, we

know the dimension of the KKT matrix must be  $(d + m) \times (d + m)$  and we have found  $d$  positive eigenvalues and  $m$  negative eigenvalues. By Sylvester's Law of Inertia, we must conclude that there are no zero eigenvalues as the total number of eigenvalues is  $d + m$ .

**2. Equality-constrained QP.** Consider an equality-constrained QP ( $G$  is symmetric)

$$\frac{1}{2}x^T Gx + c^T x \rightarrow \min \quad \text{subject to} \\ Ax = b.$$

Assume that  $A$  is full rank (i.e., its rows are linearly independent) and  $Z^T GZ$  is positive definite where  $Z$  is a basis for the null space of  $A$ , i.e.,  $AZ = 0$ .

(a) Write the KKT system for this case in matrix form.

**Solution.** Using the same process as shown in class, the KKT system for this case in matrix form is

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} -x \\ \lambda \end{bmatrix} = \begin{bmatrix} c^T \\ b \end{bmatrix}.$$

(b) Show that  $K$ , the matrix of this system, is invertible. *Hint: assume that there is a vector  $z = (x, y)^T$  such that  $Kz = 0$ . Consider the form  $z^T Kz$  and so on. You should conclude that then  $z = 0$ .*

**Solution.** We begin by writing out the assumption  $Kz = 0$ :

$$\begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0 \implies \begin{cases} Gx + A^T y = 0 \\ Ax = 0 \end{cases}.$$

Now we consider the form  $z^T Kz$  (which naturally equals 0 as a result of  $Kz$  equalling 0):

$$\begin{aligned} 0 &= \begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} G & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x^T & y^T \end{bmatrix} \begin{bmatrix} Gx + A^T y \\ Ax \end{bmatrix} = x^T Gx + x^T A^T y + y^T Ax \\ &= x^T Gx + (Ax)^T y + y^T Ax \\ &= x^T Gx \end{aligned}$$

as  $Ax = 0$  from above. Due to this,  $x \in \text{nul}\{A\}$  and hence can be written as  $x = Zv$  for some  $v \in \mathbb{R}^m$ . Therefore  $0 = x^T Gx = (Zv)^T GZv = v^T Z^T GZv$ . By positive-definiteness of  $Z^T GZ$ ,  $v = 0$  which in turn implies that  $x = 0$ . Returning to the original system, we now have

$$0 = Gx + A^T y = A^T y \implies y = 0$$

because  $A$  has full rank. Since  $x = 0$  and  $y = 0$ , we have that  $z = (x, y)^T = 0$  and therefore that  $K$  is invertible.

(c) Conclude that there exists a unique vector  $(x^*, \lambda^*)^T$  that solves the KKT system. Note that since we have only equality constraints, positivity of  $\lambda$  is irrelevant.

**Solution.** Since  $K$  is invertible, we know it is a bijection which implies that there exists a unique vector  $z$  such that

$$Kz = \begin{bmatrix} c^T \\ b \end{bmatrix}.$$

Write  $z = (x^*, \lambda^*)^T$  where  $x^* \in \mathbb{R}^d$  and  $\lambda^* \in \mathbb{R}^m$ . As such, the KKT system is solved (KKT conditions are met): the first matrix equation is the Lagrange multipliers condition and the second equation is the equality constraints, KKT conditions (i) and (ii) respectively. KKT conditions (iii) and (iv) only involve inequality constraints so they need not be met. Finally, since there are only equality constraints, condition (ii) ( $c_i(x^*) = 0$ ) implies condition (v) ( $\lambda_i^* c_i(x^*) = 0$ ). Therefore, all the conditions are met and  $z$  solves the system.