CISC 371 Class 26

KKT Geometry

Texts: [1] pp. 218–222; [2] pp. 305–317

Main Concepts:

- *Lagrange multiplier* $\lambda = 0$*: inactive constraint*
- *Lagrange multiplier* $\lambda \neq 0$*: solution is a boundary point*
- *Lagrange multiplier* $\lambda < 0$*:inconsistent constraints*
- *Equality as simultaneous inequalities*

Sample Problem, Data Analytics: How can we find a non-negative solution to a problem?

We can explore some geometrical interpretations of the KKT conditions by using a simple example. Suppose that our objective function is a quadratic form that has a size-2 vector argument, so it is defined over a Cartesian coordinate plane. We will use one or more of three linear inequalities.

26.1 Vector Example 1: Three Linear Inequalities

Our first exploration of the geometry of KKT points will use a convex optimization problem.

Example: quadratic objective, three linear inequality constraints

$$
\begin{aligned}
\vec{w}^* &= \operatorname*{argmin}_{\vec{w} \in \mathbb{R}^n} f(\vec{w}) \\
\vec{p}(\vec{w}^*) &\leq \vec{0} \\
f(\vec{w}) &= \tfrac{1}{2} \vec{w}^T \vec{w} \\
\vec{p}(\vec{w}) &= A_{3\times 2} \vec{w} - \vec{b} \\
\text{where} & \\
p_1(\vec{w}) &= [1 \;\; 0] \vec{w} - 2 \\
p_2(\vec{w}) &= [0 \;\; 1] \vec{w} + 1 \\
p_3(\vec{w}) &= [1 \;\; 1] \vec{w} + 5
\end{aligned}
\tag{26.1}
$$

We will explore four variations of Problem 26.1. These will be the objective subject to each individual constraint and the objective subject to all of the constraints. Contours, or level curves, of the objective function are shown in Figure 26.1.

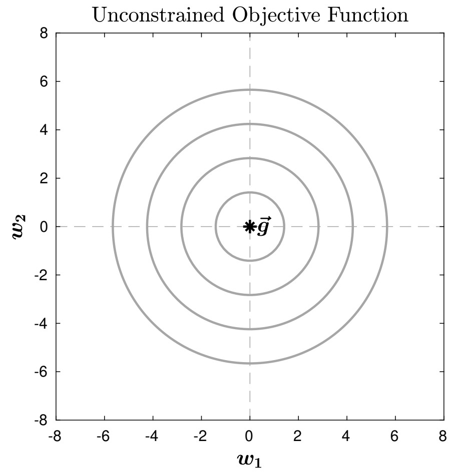

Figure 26.1: Contours of a quadratic objective function $f(\vec{w}) = \frac{1}{2}\vec{w}^T\vec{w}$.

Problem: Objective and First Linear Inequality

Our first vector problem and its Lagrange function are

$$\begin{array}{rrcl} & \vec{w}^* & = & \underset{\vec{w}\in\mathbb{R}^n}{\operatorname{argmin}} \dfrac{1}{2}\vec{w}^T\vec{w} \\ \text{where} & p_1(\vec{w}) & = & [1\ \ 0]\vec{w} - 2 \\ & \mathcal{L}_1(\vec{w}, \lambda_1) & = & \frac{1}{2}\vec{w}^T\vec{w} + \lambda_1([1\ \ 0]\vec{w} - 2) \end{array} \tag{26.2}$$

Contours of the objective function, and the level set of the first inequality constraint, are shown in Figure 26.2. Because the unconstrained minimizer is in the level set of the constraint, we know that $\lambda_1 = 0$ and that the solution to the problem is $\vec{w}_1^* = \vec{0}$.

Observation: We set $\lambda_1 = 0$ because the first constraint does not change the problem from unconstrained optimization.

Problem: Objective and Second Linear Inequality

Our second vector problem and its Lagrange function are

$$\begin{array}{rrcl} & \vec{w}^* & = & \underset{\vec{w}\in\mathbb{R}^n}{\operatorname{argmin}} \dfrac{1}{2}\vec{w}^T\vec{w} \\ \text{where} & p_2(\vec{w}) & = & [0\ \ 1]\vec{w} + 1 \\ & \mathcal{L}_2(\vec{w}, \lambda_2) & = & \frac{1}{2}\vec{w}^T\vec{w} + \lambda_2([0\ \ 1]\vec{w} + 1) \end{array} \tag{26.3}$$

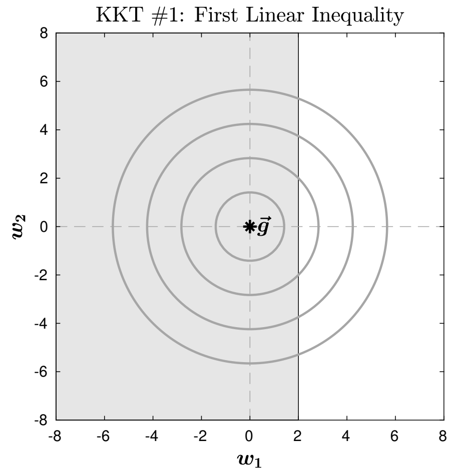

Figure 26.2: Contours of a quadratic objective function and the first linear inequality constraint. The minimizer of the unconstrained objective function is in the level set of the constraint function.

As we can see in Figure 26.3, the constraint is active. The solution to the Lagrange function in Equation 26.3 can be found geometrically or algebraically, and is

$$\vec{w}_2^* = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \qquad \lambda_2^* = +1 \tag{26.4}$$

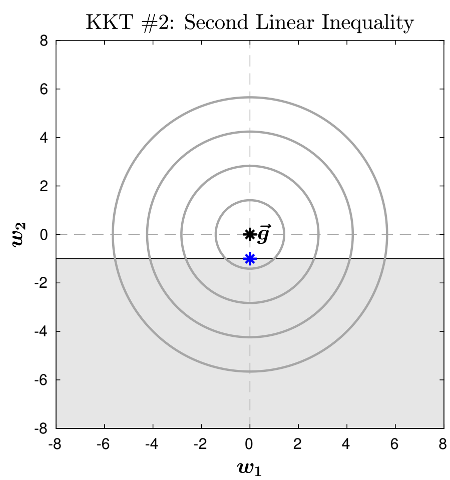

Figure 26.3: Contours of a quadratic objective function and the second linear inequality constraint. The minimizer of the unconstrained objective function is outside of the level set of the constraint function.

Problem: Objective and Third Linear Inequality

Our third vector problem and its Lagrange function are

$$\begin{array}{rrcl} & \vec{w}^* & = & \underset{\vec{w}\in\mathbb{R}^n}{\operatorname{argmin}} \frac{1}{2}\vec{w}^T\vec{w} \\ \text{where} & p_3(\vec{w}) & = & [1 \;\; 1]\vec{w} + 5 \\ & \mathcal{L}_3(\vec{w}, \lambda_3) & = & \frac{1}{2}\vec{w}^T\vec{w} + \lambda_2([1 \;\; 1]\vec{w} + 5) \end{array} \tag{26.5}$$

As we can see in Figure 26.4, this constraint is also active. The solution to the Lagrange function in Equation 26.5 can be found geometrically or algebraically, and is

$$\vec{w}_3^* = \begin{bmatrix} -1.5 \\ -1.5 \end{bmatrix} \qquad \lambda_3^* = +1.5 \tag{26.6}$$

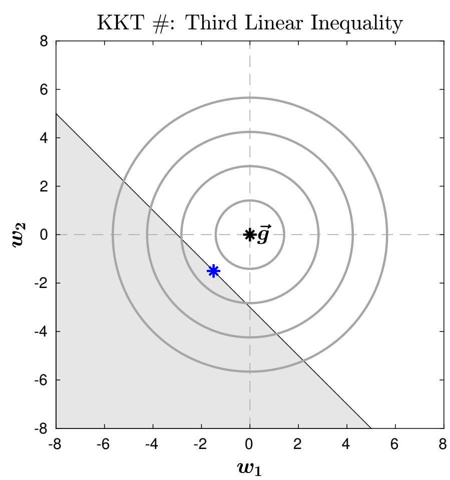

Figure 26.4: Contours of a quadratic objective function and the third linear inequality constraint. The minimizer of the unconstrained objective function is outside of the level set of the constraint function.

Problem: Objective and All Linear Inequalities

Our fourth vector problem has all three constraints. The problem and its Lagrange function are

$$\begin{aligned}
\vec{w}^* &= \underset{\vec{w}\in\mathbb{R}^n}{\operatorname{argmin}} \frac{1}{2}\vec{w}^T\vec{w} \\
\text{where} \qquad p_4(\vec{w}) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \vec{w} + \begin{bmatrix} -1 \\ +1 \\ +5 \end{bmatrix}
\end{aligned} \tag{26.7}$$

$$\mathcal{L}_4(\vec{w}, \vec{\lambda}) = \tfrac{1}{2}\vec{w}^T\vec{w} + \vec{\lambda}^T \left[\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \vec{w} + \begin{bmatrix} -1 \\ +1 \\ +5 \end{bmatrix} \right]$$

If we use a numerical method, we find that the solution to Problem 26.7 is

$$\vec{w}_4^* = \begin{bmatrix} -1.5 \\ -1.5 \end{bmatrix} \qquad \vec{\lambda} = \begin{bmatrix} 0 \\ 0 \\ +1.5 \end{bmatrix} \tag{26.8}$$

Contours of the objective function, and the feasible set,are shown in Figure 26.5.

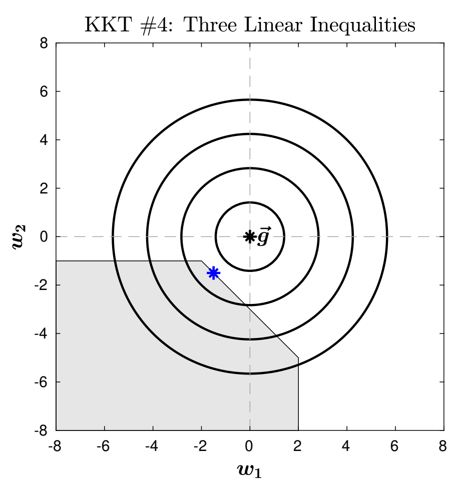

Figure 26.5: Contours of a quadratic objective function and all three linear inequality constraints. The minimizer of the unconstrained objective function is outside of the intersection the level sets of constraint functions.

Let us carefully inspect the solution in Equation 26.8. We see that the first Lagrange multiplier is zero, as it was in the solution to Problem 26.2. The geometry has not changed and this Lagrange multiplier is zero because the first constraint does not affect the basic solution to the objective function.

The second Lagrange multiplier is also zero. When we inspect Figure 26.5, we see that the second constraint is not active. This is because the third constraint provides a stationary point $\hat{w}$ that is an interior point of the level set of the second constraint. The second Lagrange multiplier is zero because the second constraint has become inactive.

The third Lagrange multiplier is positive. This is because the third constraint is active and the solution to Problem 26.7 is a boundary point of the level set of the third constraint.

Observation: At a KKT point, a Lagrange multiplier that is zero indicates that the corresponding constraint is inactive. Whether it is inactive because it does not change the unconstrained problem, or because it produces a stationary point that is an interior point of the level set of another constraint, may require careful inspection.

26.2 Vector Example 2: Two Linear Inequalities

Our second exploration of the geometry of KKT points will also use a convex optimization problem.

Example: quadratic objective, two linear inequality constraints

$$
\begin{aligned}
\vec{w}^* &= \underset{\vec{w}\in\mathbb{R}^n}{\operatorname{argmin}} f(\vec{w}) \\
\vec{p}(\vec{w}^*) &\leq \vec{0} \\
f(\vec{w}) &= \tfrac{1}{2}\vec{w}^T \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \vec{w} \\
\vec{p}(\vec{w}) &= A_{2\times 2}\vec{w} - \vec{b}
\end{aligned}
\tag{26.9}
$$

where

$$
\begin{aligned}
p_5(\vec{w}) &= [\;\;0 \;\; 1]\vec{w} + 2 \\
p_6(\vec{w}) &= [-1 \;\; 2]\vec{w} + 5
\end{aligned}
$$

We will explore three variations of Problem 26.9. These will be the objective subject to each individual constraint and the objective subject to both of the constraints. Contours, or level curves, of the objective function are shown in Figure 26.6.

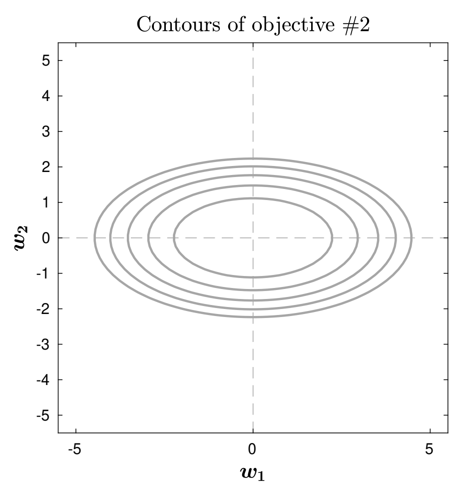

Figure 26.6: Contours of a quadratic objective function $f(\vec{w}) = w_1^2 + 4w_2^2$.

Problem: Objective and First Linear Inequality

Our next vector problem and its Lagrange function are

$$\begin{aligned} \vec{w}^* &= \underset{\vec{w}\in\mathbb{R}^n}{\operatorname{argmin}} \frac{1}{2}\vec{w}^T \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \vec{w} \\ \text{where} \qquad p_5(\vec{w}) &= [0 \;\; 1]\vec{w} + 2 \\ \mathcal{L}_5(\vec{w}, \lambda) &= \tfrac{1}{2}\vec{w}^T \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \vec{w} + \lambda([0 \;\; 1]\vec{w} + 2) \end{aligned} \tag{26.10}$$

Contours of the objective function, and the level set of the first inequality constraint, are shown in Figure 26.7.

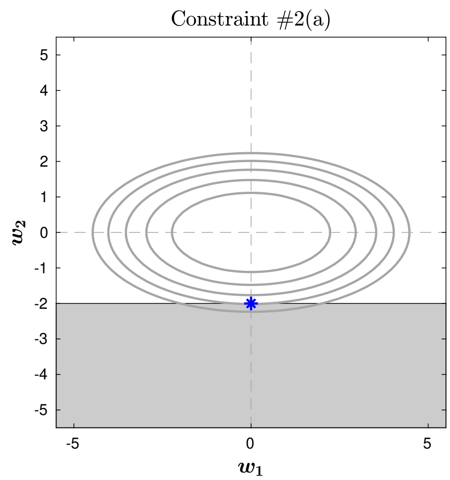

Figure 26.7: Contours of a quadratic objective function and the first linear inequality constraint.

A geometrical or numerical solution to Problem 26.10 is

$$\vec{w}_5^* = \begin{bmatrix} 0 \\ -2 \end{bmatrix} \qquad \lambda_5^* = +16 \tag{26.11}$$

The Lagrange multiplier λ_1 of Equation 26.11 is positive and, from Figure 26.7, we see that the first constraint is active.

Problem: Objective and Second Linear Inequality

Our next vector problem and its Lagrange function are

$$\begin{aligned} \vec{w}^* &= \operatorname*{argmin}_{\vec{w}\in\mathbb{R}^n} \frac{1}{2}\vec{w}^T \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \vec{w} \\ \text{where} \qquad p_6(\vec{w}) &= [0 \;\; 1]\vec{w} + 2 \\ \mathcal{L}_6(\vec{w}, \lambda_1) &= \tfrac{1}{2}\vec{w}^T \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \vec{w} + \lambda([-1 \;\; 2]\vec{w} + 5) \end{aligned} \tag{26.12}$$

Contours of the objective function, and the level set of the first inequality constraint, are shown in Figure 26.8.

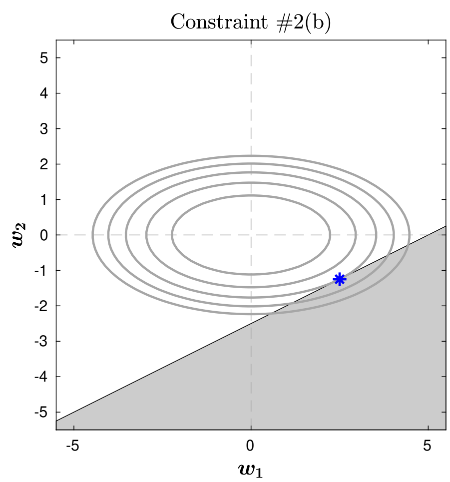

Figure 26.8: Contours of a quadratic objective function and the first linear inequality constraint.

A geometrical or numerical solution to Problem 26.10 is

$$\vec{w}_6^* = \begin{bmatrix} 2.50 \\ -1.25 \end{bmatrix} \qquad \lambda_6^* = +5 \tag{26.13}$$

The Lagrange multiplier λ_6 of Equation 26.13 is positive and, from Figure 26.8, we see that the second constraint is also active. We observe that that the second constrained solution is on a lower level curve of the objective function than was the first constrained solution.

Problem: Objective and Two Linear Inequalities

Our next vector problem and its Lagrange function are

$$\begin{aligned} \vec{w}^* &= \underset{\vec{w}\in\mathbb{R}^n}{\operatorname{argmin}} \frac{1}{2}\vec{w}^T \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \vec{w} \\ \text{where} \qquad \vec{p}(\vec{w}) &= \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \vec{w} - \begin{bmatrix} -2 \\ -5 \end{bmatrix} \\ \mathcal{L}_7(\vec{w},\vec{\lambda}) &= \tfrac{1}{2}\vec{w}^T \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \vec{w} + \vec{\lambda}^T \left[\begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \vec{w} - \begin{bmatrix} -2 \\ -5 \end{bmatrix} \right] \end{aligned} \tag{26.14}$$

Contours of the objective function, and the feasible set of the combined inequality constraints, are shown in Figure 26.9.

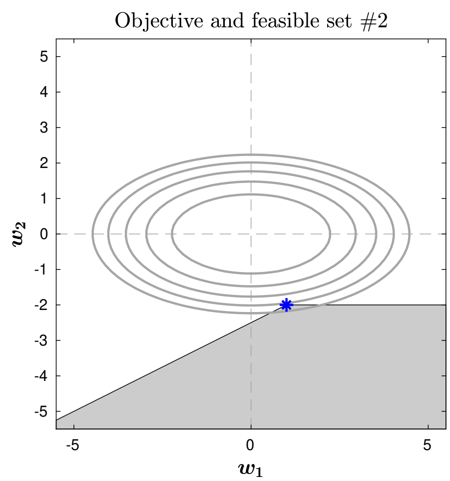

Figure 26.9: Contours of a quadratic objective function and the feasible set of two linear inequality constraints.

A geometrical or numerical solution to Problem 26.14 is

$$\vec{w}_7^* = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \qquad \vec{\lambda} = \begin{bmatrix} 12 \\ 2 \end{bmatrix} \tag{26.15}$$

The Lagrange multipliers $\vec{\lambda}$ of Equation 26.15 are both positive and, from Figure 26.9, we see that the both constraints is also active. The solution is a boundary point of the level set of each constraint, and therefore is a boundary point of the feasible set.

We observe that the Lagrange multipliers are smaller than those of the individually constrained problems. By requiring that both constraints are satisfied, the result is that a distinct minimizer is found.

26.3 Geometrical Implications of KKT Conditions

From Class 24, the conditions for a point $\hat{w}$ to be a KKT point were:

$$
\begin{array}{lrcl}
\text{(1) Primal feasibility:} & A\hat{w} & \leq & \vec{b} \\
\text{(2) Dual feasibility:} & \exists\hat{\lambda}\in\mathbb{R}^m & \geq & \vec{0} \\
\text{(3) Stationarity:} & K\hat{w}+\vec{q}+A^T\hat{\lambda} & = & \vec{0} \\
\text{(4) Complementary slackness:} & \hat{\lambda}^T\left[A\hat{w}-\vec{b}\right] & = & \vec{0}
\end{array}
\tag{26.16}
$$

Taking these conditions in turn, our geometrical observations include:

- *Primal feasibility:* a KKT point $\hat{w}$ must be in the feasible set of the problem; otherwise, the proposed solution fails to satisfy one or more constraints
- *Dual feasibility:* a Lagrange multiplier is either zero – for an inactive constraint – or is strictly positive, for an active constraint
- *Stationarity:* a KKT point $\hat{w}$ must be a minimizer of the Lagrange function, where some Lagrange multipliers may be zero at the minimizer; if every Lagrange multiplier is zero then the solution is unconstrained optimization
- *Complementary slackness:* One of two conditions must hold – either a Lagrange multiplier is zero and the constraint is inactive, or the Lagrange multiplier is positive and the KKT point $\hat{w}$ is a boundary point of the feasible set

26.4 Geometrical Interpretation of Equality Constraints

The requirement that a KKT point must have non-negative Lagrange multipliers λ_i can help us to interpret the possibly negative Lagrange multipliers that we computed with equality constraints. For example, consider a quadratic objective with a single linear equality constraint.

$$
\begin{array}{rrcl}
 & \vec{w}^* & = & \underset{\vec{w}\in\mathbb{R}^2}{\operatorname{argmin}}\,\frac{1}{2}\vec{w}^T\begin{bmatrix}2 & 0\\ 0 & 8\end{bmatrix}\vec{w} \\
\text{where} & p_1(\vec{w}) & = & \begin{bmatrix}1 & -1\end{bmatrix}\vec{w}+3 \\
 & \mathcal{L}_3(\vec{w},\mu) & = & \frac{1}{2}\vec{w}^T\begin{bmatrix}2 & 0\\ 0 & 8\end{bmatrix}\vec{w}+\mu(\begin{bmatrix}1 & -1\end{bmatrix}\vec{w}+3)
\end{array}
\tag{26.17}
$$

A geometrical or numerical solution to Problem 26.17 is

$$
\vec{w}^*=\begin{bmatrix}2.4\\ -0.6\end{bmatrix} \qquad \mu^*=-4.8 \tag{26.18}
$$

With our understanding of inequality constraints, we can see that any equality constraint can be written as two inequality constraints that must both be satisfied: one in the "positive" form and one in the "negative" form. A single linear equality constraint that is

$$\underline{m}\hat{w} - c = 0 \tag{26.19}$$

is equivalent to the two inequality constraints

$$\begin{array}{rcl} \underline{m}\hat{w} - c & \leq & 0 \\ -\underline{m}\hat{w} + c & \leq & 0 \end{array} \tag{26.20}$$

One of the constraints in Equation 26.20 would have a Lagrange multiplier of zero and the other constraint would have a positive Lagrange multiplier. Numerically we cannot determine which constraint the user intended, and the result is that we might calculate a Lagrange multiplier $\mu < 0$. Figure 26.10 illustrates the level sets, with transposes of the gradients, for the constraint in Problem 26.17 and the sign reversal described in Equation 26.20.

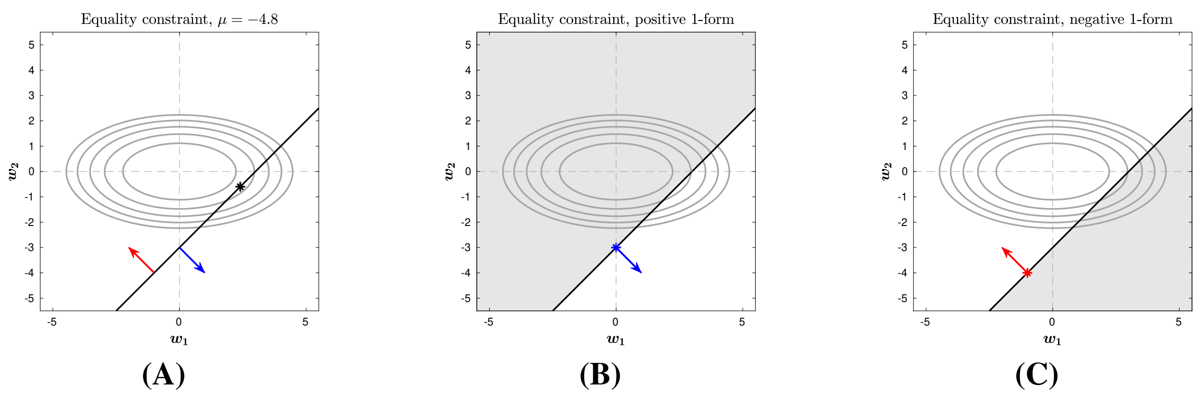

Figure 26.10: Contours of a quadratic objective function, in gray, and interpretations of a linear equality constraint. (A) An equality constraint and its solution, in black, with possible transposed gradients, as blue and red arrows. (B) One side of a corresponding inequality constraint, with the objective minimizer as an interior point of the level set of the constraint. (C) Second side of a corresponding inequality constraint, with the objective minimizer outside of the level set of the constraint.

References

[1] Beck A: Introduction to Nonlinear Optimization: Theory, Algorithms, and Applications with MATLAB. Siam Press, 2014

[2] Boyd S, Vandenberghe L: Convex Optimization. Cambridge University Press, 2004