Chapter 1

Simple equality constrained optimization

1.1 Convexity

The problem is not convex, since $\Omega = \left\{\vec{\boldsymbol{x}} \,\middle|\, x_1^2 + x_2^2 - 1 = 0\right\}$ is not convex.
Counterexample: $\vec{\boldsymbol{a}} := \begin{bmatrix}1\\0\end{bmatrix} \in \Omega$ and $\vec{\boldsymbol{b}} := \begin{bmatrix}-1\\0\end{bmatrix} \in \Omega$ but $\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}} = \vec{\boldsymbol{0}} \notin \Omega$

1.2 Lagrangian

$$\mathcal{L}\left(x_1, x_2, \lambda\right) = f(x) - \lambda g(x) = x_2 - \lambda\left(x_1^2 + x_2^2 - 1\right)$$

1.3 First Order Necessary Conditions

If LICQ holds at $\vec{\boldsymbol{x}}^*$ (which it does for every $\vec{\boldsymbol{x}}^* \in \Omega$ in this case) and $\vec{\boldsymbol{x}}^*$ is a local minimizer then there exists a λ^* such that $\nabla_x \mathcal{L}(\vec{\boldsymbol{x}}^*, \lambda^*) = 0$ and $\nabla_\lambda \mathcal{L}(\vec{\boldsymbol{x}}^*, \lambda^*) = 0$ where:

$$\begin{aligned}
\nabla_x \mathcal{L}(\vec{\boldsymbol{x}}, \lambda) &= \begin{bmatrix}0\\1\end{bmatrix} - \lambda \begin{bmatrix}2x_1\\2x_2\end{bmatrix} \overset{!}{=} \vec{\boldsymbol{0}} \\
\nabla_\lambda \mathcal{L}(\vec{\boldsymbol{x}}, \lambda) &= x_1^2 + x_2^2 - 1
\end{aligned}$$

1.4 Stationary Points

System to solve:

$$\begin{aligned}
-2\lambda x_1 &= 0 && \implies \lambda = 0 \lor x_1 = 0 \\
-2\lambda x_2 + 1 &= 0 \\
x_1^2 + x_2^2 - 1 &= 0
\end{aligned}$$

Assuming $\lambda = 0$ then $-2\lambda x_2 + 1 = 0 \implies 1 = 0$ is a contradiction, so it follows that $x_1 = 0$. From $x_1 = 0$ it follows that $x_2^2 + 0 = 1$ and therefore $x_{2,a} = -1$, $x_{2,b} = 1$. From the second equation in the system we derive:

$$\begin{aligned}
-2\lambda x_2 + 1 &= 0 \\
2\lambda x_2 &= 1 \\
\lambda &= \frac{1}{2x_2}
\end{aligned}$$

so for $x_2 = \pm 1$ we have $\lambda = \mp\frac{1}{2}$ and the stationary points are:

$$\begin{aligned}
\vec{\boldsymbol{a}} &:= \begin{bmatrix}0\\-1\end{bmatrix} \quad \lambda_a = -\frac{1}{2} \\
\vec{\boldsymbol{b}} &:= \begin{bmatrix}0\\1\end{bmatrix} \quad \lambda_b = \frac{1}{2}
\end{aligned}$$

1.5 Implication for global minimizer

Since the global minimizer exists in this case and has to be one of the two stationary points where FONC hold, namely the one that minimizes x_2, it is easy to see that the global minimizer is $\vec{\boldsymbol{x}}^* = \vec{\boldsymbol{a}} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

1.6 Second order necessary conditions

At $\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}$ FONC holds, so for the second order necessary conditions it remains to show that:

$$Z^T \nabla_x^2 \mathcal{L}(\vec{\boldsymbol{x}}^*, \lambda^*) Z \geq 0$$

with:

$$\begin{aligned} \nabla_x \mathcal{L}(\vec{\boldsymbol{x}}, \lambda) &= \begin{bmatrix} -2\lambda x_1 \\ -2\lambda x_2 + 1 \end{bmatrix} \\ \nabla_x^2 \mathcal{L}(\vec{\boldsymbol{x}}, \lambda) &= \begin{bmatrix} -2\lambda & 0 \\ 0 & -2\lambda \end{bmatrix} = -2\lambda \mathbb{1} \\ \nabla g(x)^T &= \begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix} \end{aligned}$$

and since Z spans the null-space of $\nabla g(x)^T$:

$$\begin{aligned} &\begin{bmatrix} 2x_1 & 2x_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 0 = 2x_1 z_1 + 2x_2 z_2 \\ &\overset{\text{e.g.}}{\Longrightarrow} Z = \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix} \end{aligned}$$

so the SONC read, at the stationary point $\vec{\boldsymbol{a}}$:

$$\begin{aligned} Z_a^T \nabla_x^2 \mathcal{L}\left(\vec{\boldsymbol{a}}, -\frac{1}{2}\right) Z_a &= Z_a^T (-2\lambda_a \mathbb{1}) Z_a \\ &= -2\lambda_a \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= -2\lambda_a = 1 \\ &\geq 0 \end{aligned}$$

which are fulfilled, while at stationary point $\vec{\boldsymbol{b}}$:

$$\begin{aligned} Z_b^T \nabla_x^2 \mathcal{L}\left(\vec{\boldsymbol{b}}, \frac{1}{2}\right) Z_b &= Z_b^T (-2\lambda_b \mathbb{1}) Z_b \\ &= -2\lambda_b = -1 \\ &< 0 \end{aligned}$$

the SONC are not fulfilled.

1.7 Additional Constraint violating LICQ

For the stationary point $\vec{\boldsymbol{a}}$ add a constraint g_2:

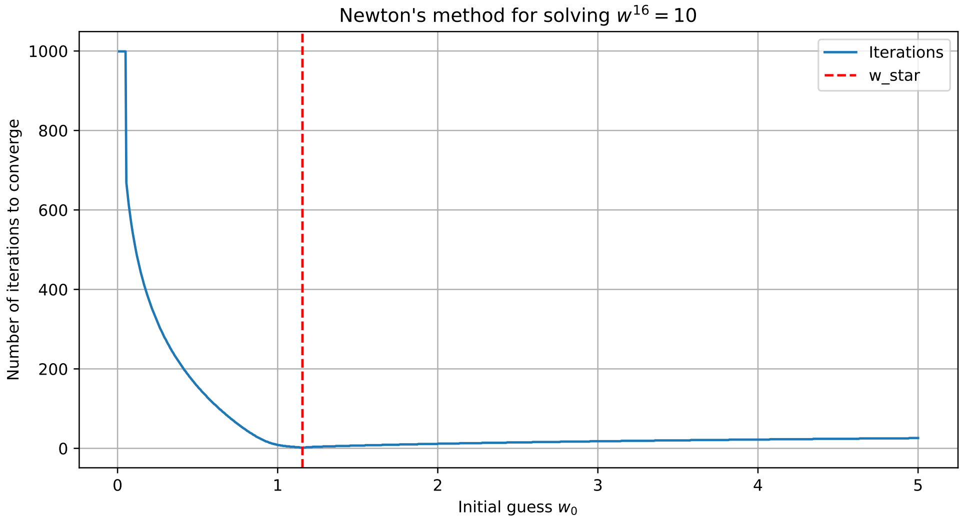
$$\begin{aligned} g_1(\vec{\boldsymbol{x}}) &= x_1^2 + x_2^2 - 1 \overset{!}{=} 0 \\ g_2(\vec{\boldsymbol{x}}) &= x_2 + 1 \overset{!}{=} 0 \\ \nabla_x g_1(\vec{\boldsymbol{x}}) &= \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} \\ \nabla_x g_2(\vec{\boldsymbol{x}}) &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

at $\vec{\boldsymbol{a}} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \implies \nabla_x g_1(\vec{\boldsymbol{a}}) = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$ the two gradients are not linearly independent, since:

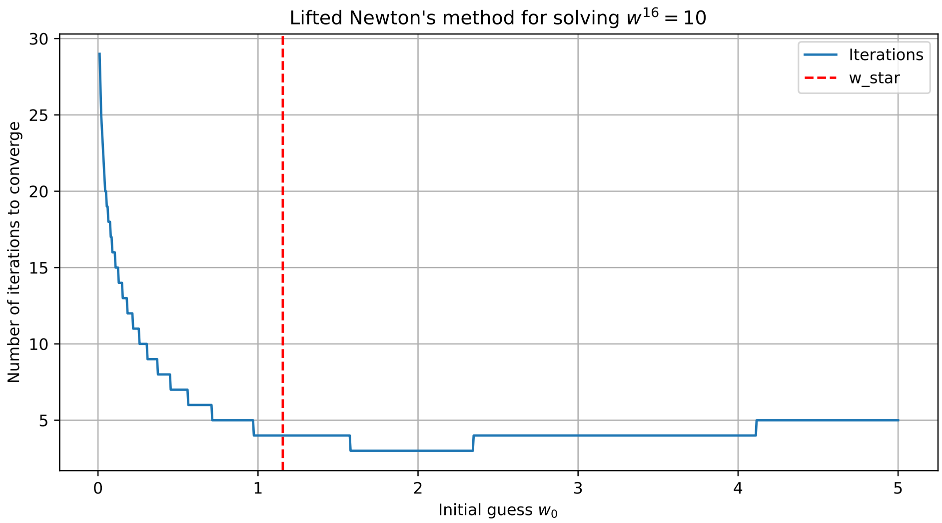
$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \nabla_x g_2(\vec{\boldsymbol{a}}) = -2\nabla_x g_1(\vec{\boldsymbol{a}}) = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

Chapter 2

Lifted Newton method



(a) Iteration count over initial guess for regular Newton's method



(b) Iteration count over initial guess for lifted Newton's method

Figure 2.1: The comparison of the lifted-and non-lifted approach reveals that while the distribution of iterations over initial guesses appears similar in both approaches, the iteration count for the lifted method is consistently more efficient by orders of magnitude. The regular method appears to have its lowest iteration count where the initial guess and solution coincide, while the lifted method interestingly converges faster for an initial guess around $w_0 = 2$, or slightly above the solution $w^* = \sqrt[16]{10}$. While both methods struggle for low initial guesses, where progress can only be made at an excruciatingly slow pace due to the exponent of 15 in the function gradient, the lifted method appears to be much more robust to such averse choices of initial value.