CE 191: Civil and Environmental Engineering Systems Analysis

LEC 05 : Optimality Conditions

Professor Scott Moura
Civil & Environmental Engineering
University of California, Berkeley

Fall 2014

Conditions for Optimality

Consider an unconstrained QP

$$\min \qquad f(x) = \frac{1}{2} x^T Q x + R x$$

Recall from calculus (e.g. Math 1A) the first order necessary condition (FONC) for optimality: If x^* is an optimum, then it must satisfy

$$\begin{aligned} \frac{d}{dx} f(x^*) &= 0 \\ &= Qx^* + R = 0 \\ &\Rightarrow \boxed{x^* = -Q^{-1} R} \end{aligned}$$

Also recall the second order sufficiency condition (SOSC): If $x^\dagger$ is a stationary point (i.e. it satisfies the FONC), then it is also a minimum if

$$\begin{aligned} \frac{\partial^2}{\partial x^2} f(x^\dagger) & \qquad \text{positive definite} \\ \Rightarrow Q & \qquad \text{positive definite} \end{aligned}$$

Review: Positive-definite matrices

All of the following conditions are equivalent:

Consider $Q \in \mathbb{R}^{n \times n}$

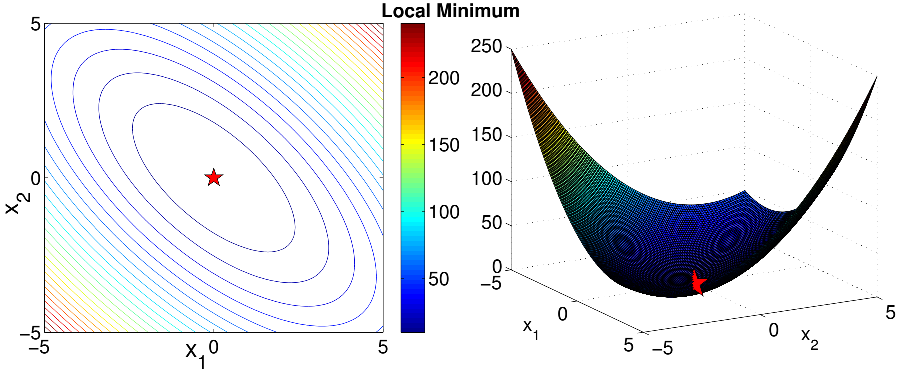
- Q is positive definite
- $x^T Q x > 0, \forall x \neq 0$
- the real parts of all eigenvalues of Q are positive
- $-Q$ is negative definite

- Q is positive semi-definite
- $x^T Q x \geq 0, \forall x \neq 0$
- the real parts of all eigenvalues of Q are positive, and at least one eigenvalue is zero
- $-Q$ is negative semi-definite

Nature of stationary point based on SOSC

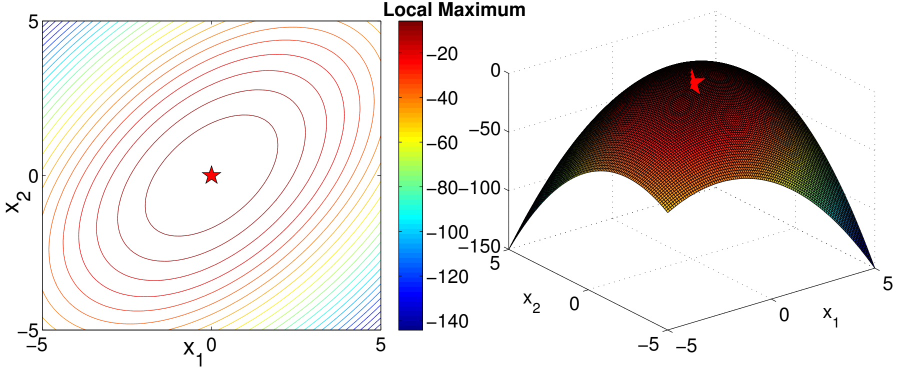
Hessian matrix	Quadratic form	Nature of $x^\dagger$
positive definite	$x^T Q x > 0$	local minimum
negative definite	$x^T Q x < 0$	local maximum
positive semi-definite	$x^T Q x \geq 0$	valley
negative semi-definite	$x^T Q x \leq 0$	ridge
indefinite	$x^T Q x$ any sign	saddle point

Local Minimum



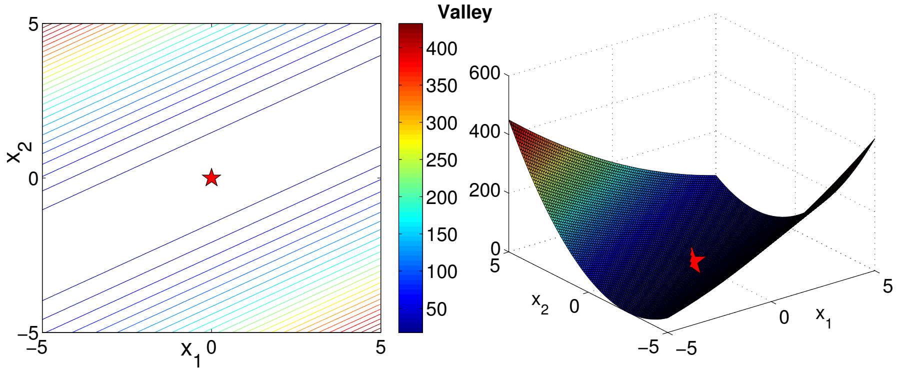
$$Q = \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix}, \qquad \text{eig}(Q) = \{1, 5\}$$

Local Maximum



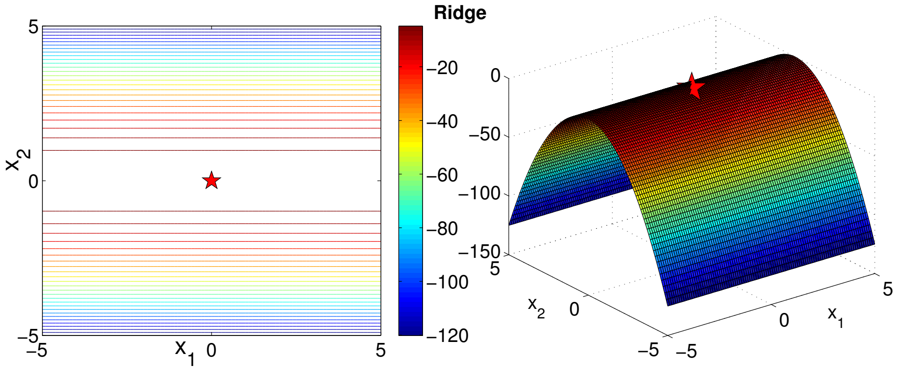
$$Q = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}, \qquad \text{eig}(Q) = \{-3, -1\}$$

Valley



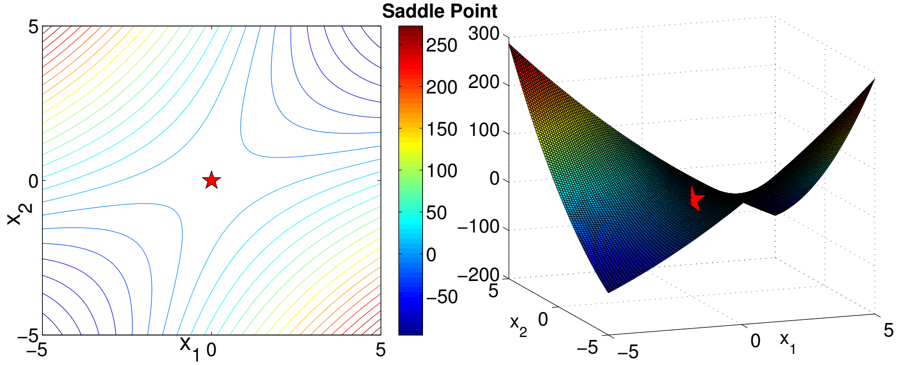
$$Q = \begin{bmatrix} 8 & -4 \\ -4 & 2 \end{bmatrix}, \qquad \text{eig}(Q) = \{0, 10\}$$

Ridge



$$Q = \begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathrm{eig}(Q) = \{-5, 0\}$$

Saddle Point



$$Q = \begin{bmatrix} 2 & -4 \\ -4 & 1.5 \end{bmatrix}, \qquad \text{eig}(Q) = \{-2.26, 7.76\}$$

Example 5.1

$$
\begin{aligned}
f(x_1, x_2) &= (3 - x_1)^2 + (4 - x_2)^2 \\
&= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -6 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\end{aligned}
$$

Check the FONC:

$$
\frac{\partial f}{\partial x} = \begin{bmatrix} -6 + 2x_1 \\ -8 + 2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
$$

has the solution $(x_1^\dagger, x_2^\dagger) = (3, 4)$.

Check the SOFC:

$$
\frac{\partial^2 f}{\partial x^2} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow \text{positive definite}
$$

<u>Solution:</u> Unique local minimum

Example 5.2

$$
\begin{aligned}
f(x_1, x_2) &= -4x_1 + 2x_2 + 4x_1^2 - 4x_1x_2 + x_2^2 \\
&= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} -4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
\end{aligned}
$$

Check the FONC:

$$
\frac{\partial f}{\partial x} = \begin{bmatrix} -4 + 8x_1 - 4x_2 \\ 2 - 4x_1 + 2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
$$

has an infinity of solutions $(x_1^\dagger, x_2^\dagger)$ on the line $2x_1 - x_2 = 1$.

Check the SOFC:

$$
\frac{\partial^2 f}{\partial x^2} = \begin{bmatrix} 8 & -4 \\ -4 & 2 \end{bmatrix} \rightarrow \text{positive semidefinite}
$$

Solution: Infinite set of minima (valley)

A Connection to Nonlinear Programming

Consider the more general nonlinear programming problem, with nonlinear cost and nonlinear constraints

$$\begin{aligned} \min \quad & J = f(x) \\ \text{s. to} \quad & g(x) \leq 0 \end{aligned}$$

Consider a given value for the decision variable, x_k.
Let's take a Taylor series expansion of the cost and constraints.

Taylor Series

Review: Expand $f(x)$ into infinite power series around $x = x_k$

$$\begin{aligned} f(x) &= f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2 + \ldots \\ &= \sum_{n=0}^{\infty} \frac{f^n(x_k)}{n!} (x - x_k)^n \end{aligned}$$

Expand cost function, truncated to be 2nd order

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k) + \frac{1}{2} f''(x_k)(x - x_k)^2$$

Expand inequality constraints, truncated to be 1st order

$$g(x) \approx g(x_k) + g'(x_k)(x - x_k) \leq 0$$

For ease of notation, define $\tilde{x} = x - x_k$

Sequential Quadratic Programming (SQP)

We arrive at the following approximate QP

$$\begin{aligned} \min \quad & Q\tilde{x}^2 + R\tilde{x} \\ \text{s. to} \quad & A\tilde{x} \leq b \end{aligned}$$

where

$$\begin{aligned} Q &= \frac{1}{2}f''(x_k), \qquad R = f'(x_k) \\ A &= g'(x_k), \qquad b = -g(x_k) \end{aligned}$$

Suppose the optimal solution is $\tilde{x}^*$.
Then let $x_{k+1} = x_k + \tilde{x}^*$.
Repeat.

SQP Remarks

Remark 1:

Can add equality constraints $h(x) = 0$ and expand via 1st order Taylor series.

Remark 2:

If x_{k+1} does not satisfy $g(x_{k+1}) \leq 0$,

then you can "project" x_{k+1} onto surface $g(\cdot) = 0$.

Remark 3:

Iterate until a stopping criteria is reached, e.g. $\tilde{x} \leq \varepsilon$.

Summary:

Can re-formulate nonlinear program into sequence of quadratic programs.

Example 5.3

Consider the NLP

$$\begin{aligned} \min_{x_1, x_2} \quad & e^{-x_1} + (x_2 - 2)^2 \\ \text{s. to} \quad & x_1 x_2 \leq 1. \end{aligned}$$

with the initial guess $[x_{1,0}, x_{2,0}]^T = [1, 1]^T$. Perform 3 iterations of SQP.

We have the functions:

$f(x) = e^{-x_1} + (x_2 - 2)^2$ and $g(x) = x_1 x_2 - 1$

We seek to find the approximate QP subproblem

$$\begin{aligned} \min \quad & \frac{1}{2}\tilde{x}^T Q \tilde{x} + R^T \tilde{x} \\ \text{s. to} \quad & A\tilde{x} \leq b \end{aligned}$$

Example 5.3

Taking derivatives of $f(x)$ and $g(x)$,

$$\begin{aligned} Q &= \begin{bmatrix} e^{-x_1} & 0 \\ 0 & 2 \end{bmatrix}, \qquad R = \begin{bmatrix} -e^{-x_1} \\ 2(x_2 - 2) \end{bmatrix}, \\ A &= \begin{bmatrix} x_2 & x_1 \end{bmatrix}, \qquad b = 1 - x_1 x_2 \end{aligned}$$

Now consider the initial guess $[x_{1,0}, x_{2,0}]^T = [1, 1]^T$. This iterate is feasible.

First iteration: Q, R, A, b matrices are

$$\begin{aligned} Q &= \begin{bmatrix} e^{-1} & 0 \\ 0 & 2 \end{bmatrix}, \qquad R = \begin{bmatrix} -e^{-1} \\ -2 \end{bmatrix}, \\ A &= \begin{bmatrix} 1 & 1 \end{bmatrix}, \qquad b = 0 \end{aligned}$$

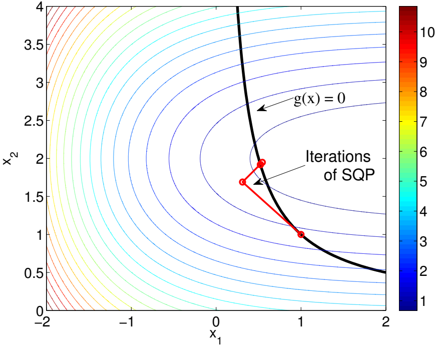
Solving this QP subproblem results in $\tilde{x}^* = [-0.6893, 0.6893]$.

Next iterate: $[x_{1,1}, x_{2,1}] = [x_{1,0}, x_{2,0}] + \tilde{x}^* = [0.3107, 1.6893]$
Iterate is feasible.

Example 5.3

Second iteration: Result is $[x_{1,2}, x_{2,2}] = [0.5443, 1.9483]$. Iterate is infeasible.

Continuing the process...



Iter.	$[x_1, x_2]$	$f(x)$	$g(x)$
0	[1, 1]	1.3679	0
1	[0.3107, 1.6893]	0.8295	-0.4751
2	[0.5443, 1.9483]	0.5829	0.0605
3	[0.5220, 1.9171]	0.6002	0.0001
4	[0.5211, 1.9192]	0.6004	-0.0000

Additional Reading

- Papalambros & Wilde Section 4.2 - Local Approximations
- Papalambros & Wilde Section 4.3 - Optimality Conditions
- Papalambros & Wilde Section 7.7 - Sequential Quadratic Programming