Optimization with constraints

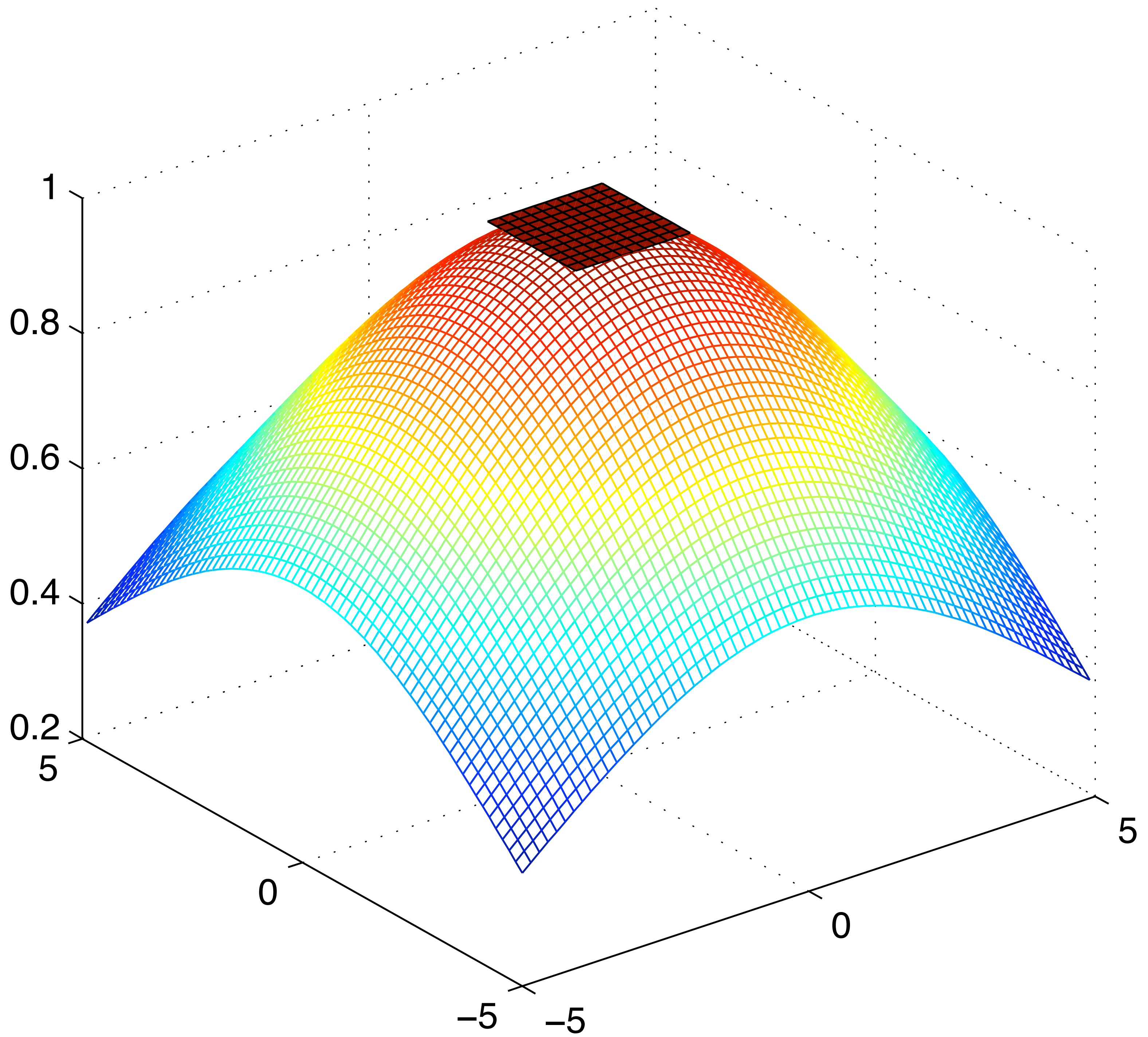
Optimization With Constraints

- Let $f(\mathbf{x})$: $R^n \rightarrow R$ (a cost function)
- Let $\mathbf{g}(\mathbf{x})$: $R^n \rightarrow R^n$
- Solve

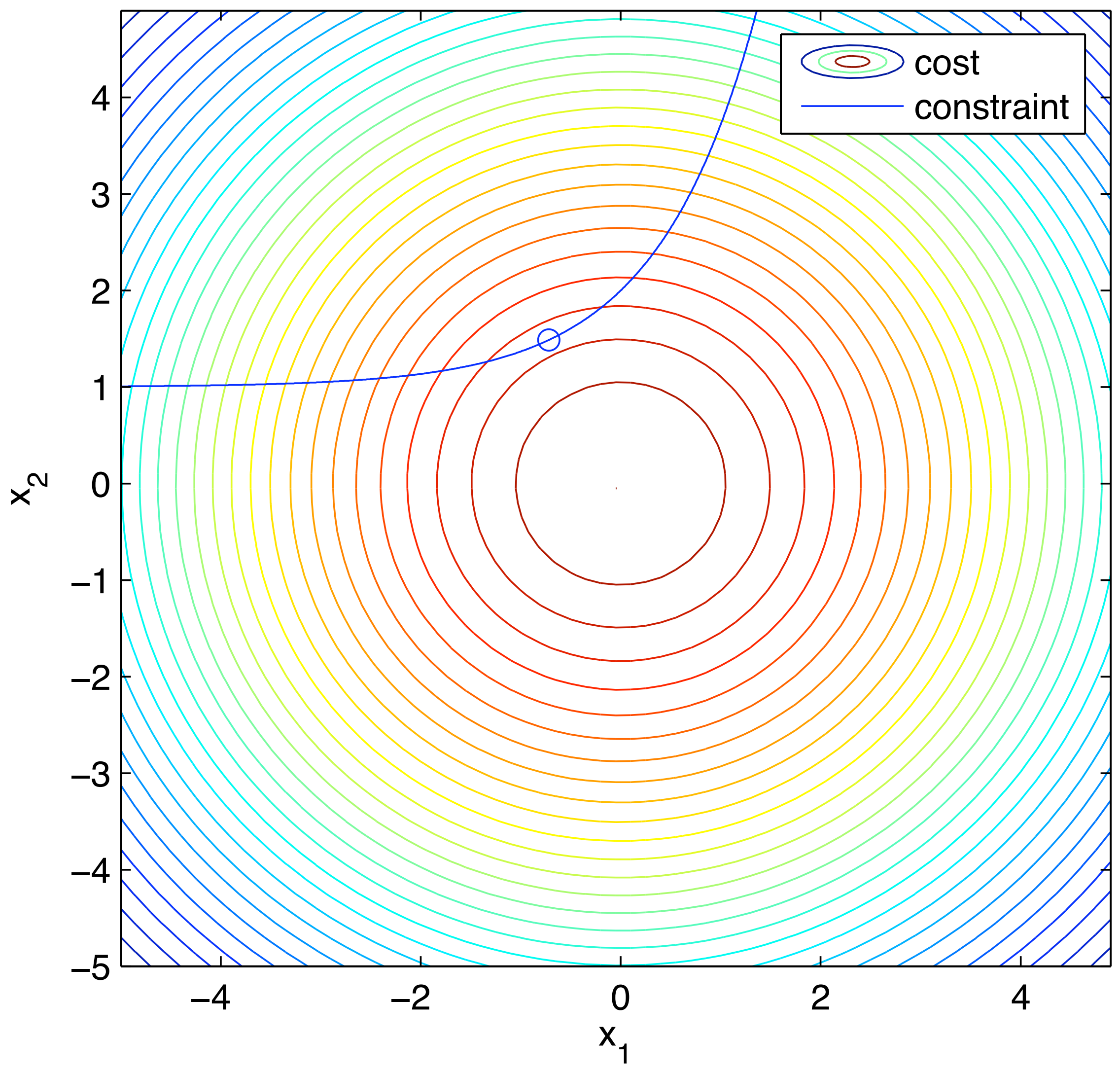
$$\min_{\mathbf{x}} \quad f(\mathbf{x})$$
$$s.t. \quad \mathbf{g}(\mathbf{x}) = \mathbf{0}$$

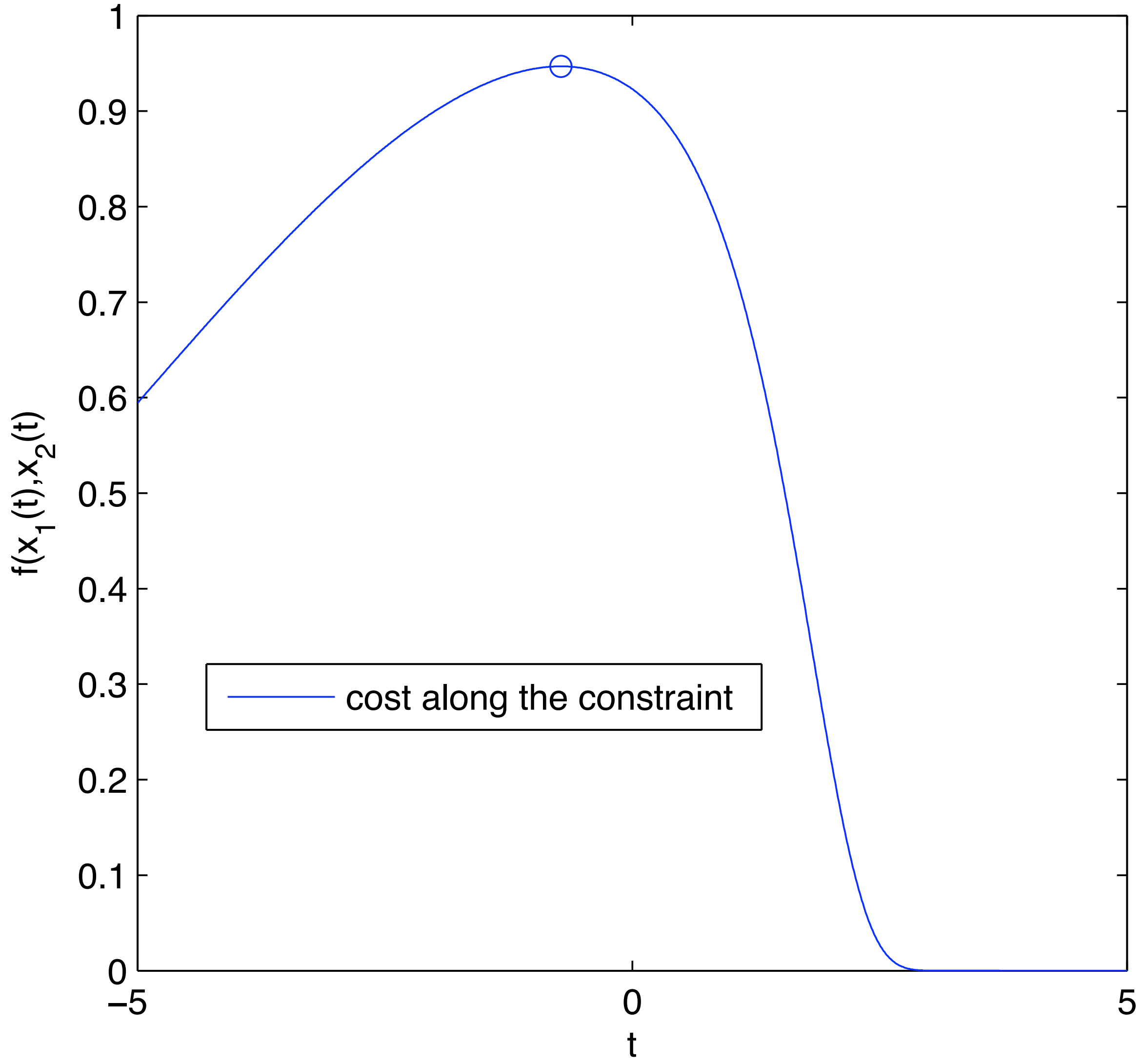
- f can be the optimization criterion and $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ can define the dynamics of the system, i.e. what possible values $\mathbf{x}$ can take within the physical system constraints.

- Without constraints: $$\left[\frac{\partial f}{\partial \mathbf{x}}\right]^T\bigg|_{\mathbf{x}^*} = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]^T\bigg|_{\mathbf{x}^*} = \mathbf{0}$$



- With constraints:

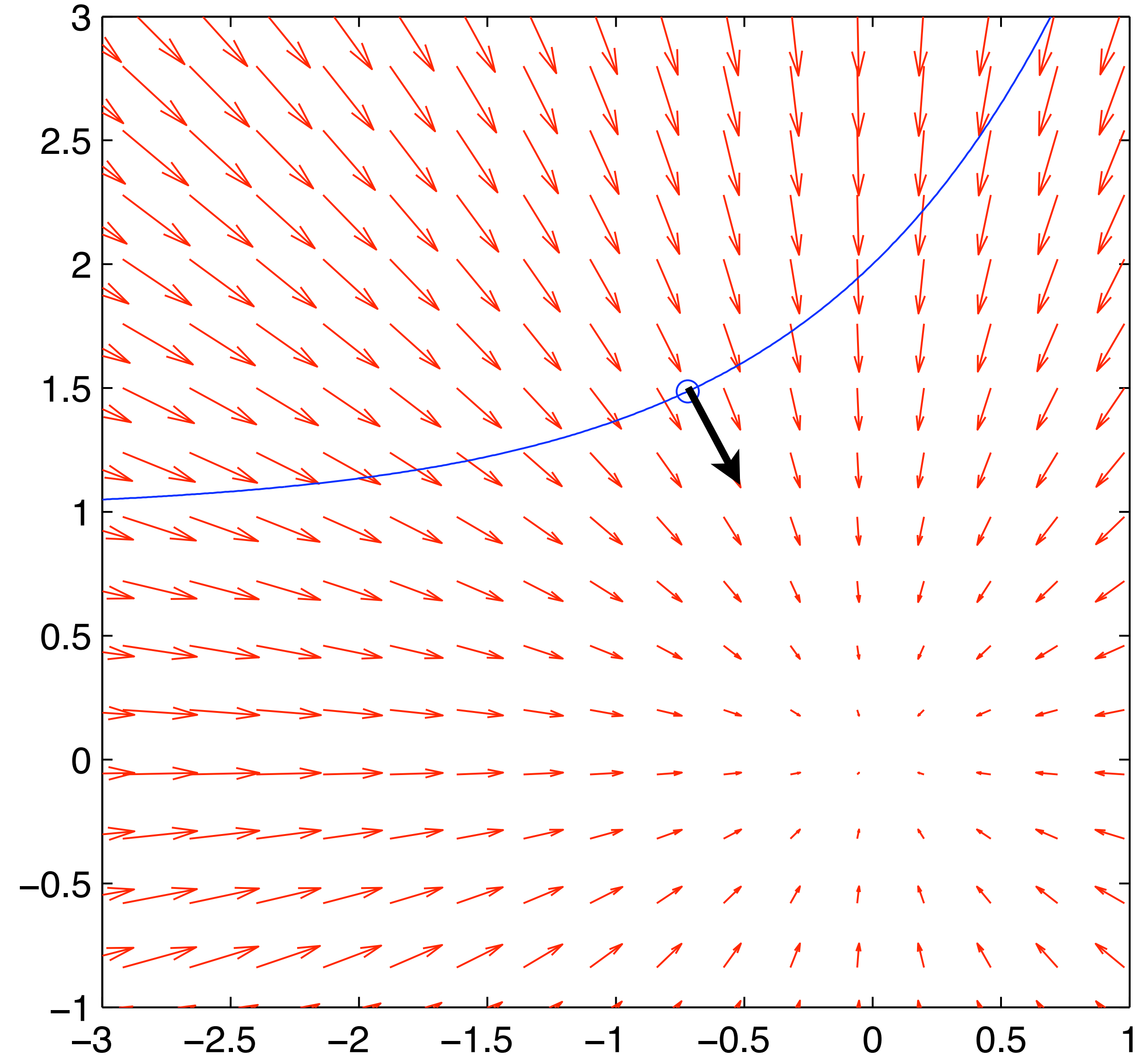




- Note that $g_i(\mathbf{x}) = 0$ defines a specific height line of the function
- Therefore, $\partial g_i / \partial \mathbf{x}$ must be perpendicular to the line
- **Claim**: If $\mathbf{x}_0$ is an extremum of $f(\mathbf{x})$ under scalar constraint $g(\mathbf{x}) = 0$ then there exists a constant $\lambda \in R$ such that

$$\left.\frac{\partial f}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} = \lambda \left.\frac{\partial g}{\partial \mathbf{x}}\right|_{\mathbf{x}_0}$$

- A small movement along the constraint line is perpendicular to the cost function gradient and therefore will not decrease f



- **Claim**: If $\mathbf{x}_0$ is an extremum of $f(\mathbf{x})$ under constraint $\boldsymbol{g}(\mathbf{x}) = 0$ then there exists a constant vector $\boldsymbol{\lambda}^n \in R^n$ such that

$$\left.\frac{\partial f}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} = \boldsymbol{\lambda}^T \left.\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right|_{\mathbf{x}_0}$$

  where

$$\left.\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right|_{\mathbf{x}_0} = \begin{bmatrix} \left.\frac{\partial g_1}{\partial x_1}\right|_{\mathbf{x}_0} & \cdots & \left.\frac{\partial g_1}{\partial x_n}\right|_{\mathbf{x}_0} \\ \vdots & & \vdots \\ \left.\frac{\partial g_m}{\partial x_1}\right|_{\mathbf{x}_0} & \cdots & \left.\frac{\partial g_m}{\partial x_n}\right|_{\mathbf{x}_0} \end{bmatrix}$$

  is the Jacobian.

- So at the constrained extremum, if there is any possible direction of change in **x** that obeys the constraints, it is in the null-space of the Jacobian and, therefore, will not change the value of *f*.

- Define the *Lagrangian* to be

$$L = f(\mathbf{x}) - \boldsymbol{\lambda}^T \mathbf{g}(\mathbf{x})$$

- Clearly

$$\frac{\partial L}{\partial \mathbf{x}} = \frac{\partial f}{\partial \mathbf{x}} - \boldsymbol{\lambda}^T \frac{\partial \mathbf{g}}{\partial \mathbf{x}}$$

- Also

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = -\mathbf{g}(\mathbf{x})$$

- So, given a pair $(\mathbf{x}, \boldsymbol{\lambda})$ such that $\frac{\partial L}{\partial \mathbf{x}} = 0$ and $\frac{\partial L}{\partial \boldsymbol{\lambda}} = 0$ then for this pair $\frac{\partial f}{\partial \mathbf{x}} = \boldsymbol{\lambda}^T \frac{\partial \mathbf{g}}{\partial \mathbf{x}}$ and $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ which means that $\mathbf{x}$ is a constrained extremum of f

Example: moving in minimal time + energy

- Mass on a frictionless track. Rests at position 0 at time 0.
- Apply constant force, k on the mass till it reaches a given position s.
- Let t be the time it takes the mass to reach s.
- Find the force k and time t_f that minimize the cost:

$$f(k, t) = ks + rt$$

  where r is a constant that defines the tradeoff between minimal time and minimal energy

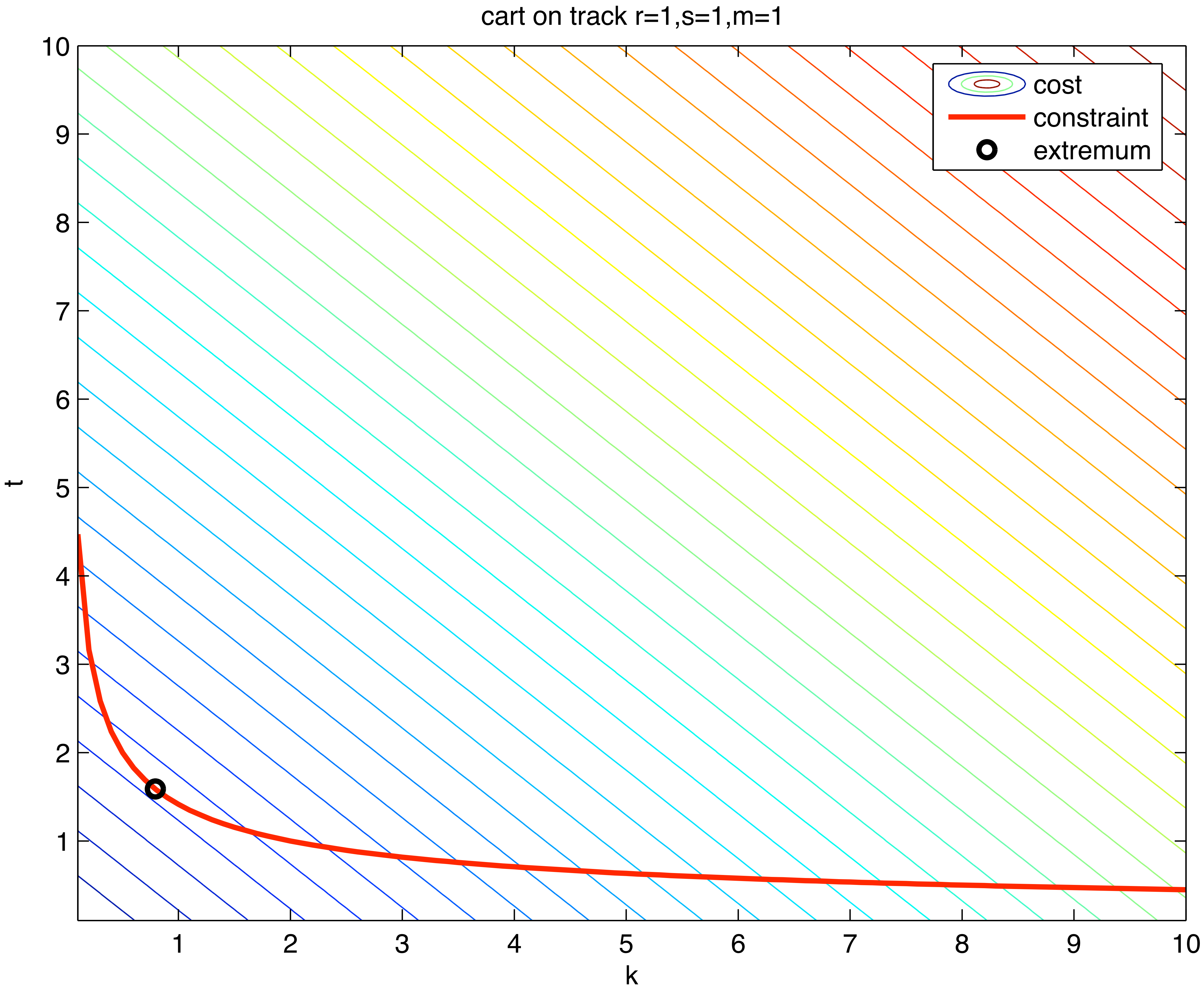
- Note that $\mathbf{x} = (t,k)$ is a solution vector whose variables are constrained by system's physics:

$$\ddot{y} = \frac{k}{m}$$

$$\dot{y}(t) = \int_0^t \frac{k}{m} d\tau = \frac{k}{m} t$$

$$y(t) = \int_0^t \frac{k}{m} t d\tau = \frac{kt^2}{2m}$$

$$g(k, t) = \frac{kt^2}{2m} - s = 0$$


cart on track r=1,s=1,m=1
cost
constraint
extremum
t
k
1
2
3
4
5
6
7
8
9
10

על מנת למצוא זוג (k,t) שממזער את f תחת אילוץ המערכת נגדיר את הLagrangian

$$L = f(k,t) - \lambda g(k,t) = ks + rt - \lambda\left(\frac{kt^2}{2m} - s\right)$$

נגזור לפי k, t, וλ לקבלת 3 משוואות: .

$$\begin{aligned}
\frac{\partial L}{\partial k} &= s - \frac{\lambda t^2}{2m} = 0 \\
\frac{\partial L}{\partial t} &= r - \frac{\lambda k t}{m} = 0 \\
\frac{\partial L}{\partial \lambda} &= \frac{kt^2}{2m} - s = 0
\end{aligned}$$

נשים לב שהמשוואה האחרונה פשוט החזירה את האילוץ. כעת ניתן לפתור עבור λ, להציב את הפתרון חזרה במשוואות ולפתור עבור (k,t)

$$\begin{aligned}\frac{\partial L}{\partial k} &= s - \frac{\lambda t^2}{2m} = 0\\ \frac{\partial L}{\partial t} &= r - \frac{\lambda k t}{m} = 0\\ \frac{\partial L}{\partial \lambda} &= \frac{kt^2}{2m} - s = 0\end{aligned}$$

מהמשוואה השנייה

$$\lambda = \frac{mr}{kt}$$

מהצבת λ במשוואה הראשונה

$$\begin{aligned}s - \frac{rt}{2k} &= 0\\ k &= \frac{rt}{2s}\end{aligned}$$

מהצבת k במשוואה השלישית נמצא ש $t = \left(\frac{4ms^2}{r}\right)^{\frac{1}{3}}$ ואם כעת נציב את t במשוואה הראשונה נקבל $k = \left(\frac{r^2 m}{2s}\right)^{\frac{1}{3}}$