Optimality Criteria

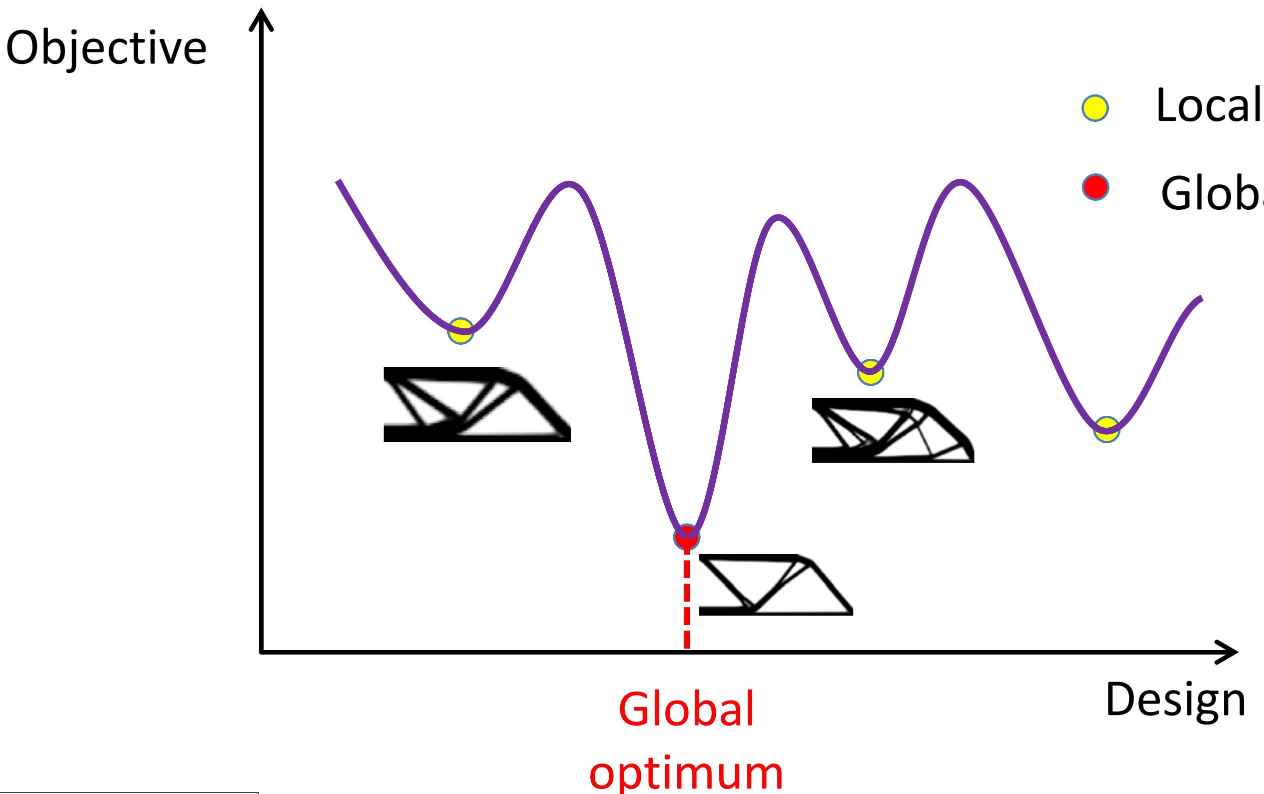
Unconstrained Optimization

How do I know the current design is an optimum

Optimum design should satisfy optimality conditions

GLOBAL VERSUS LOCAL MINIMUM

- Optimization algorithm searches for local minimum…global minimum is not guaranteed
- Starting with different initial designs will result in different designs



GLOBAL OPTIMIZATION

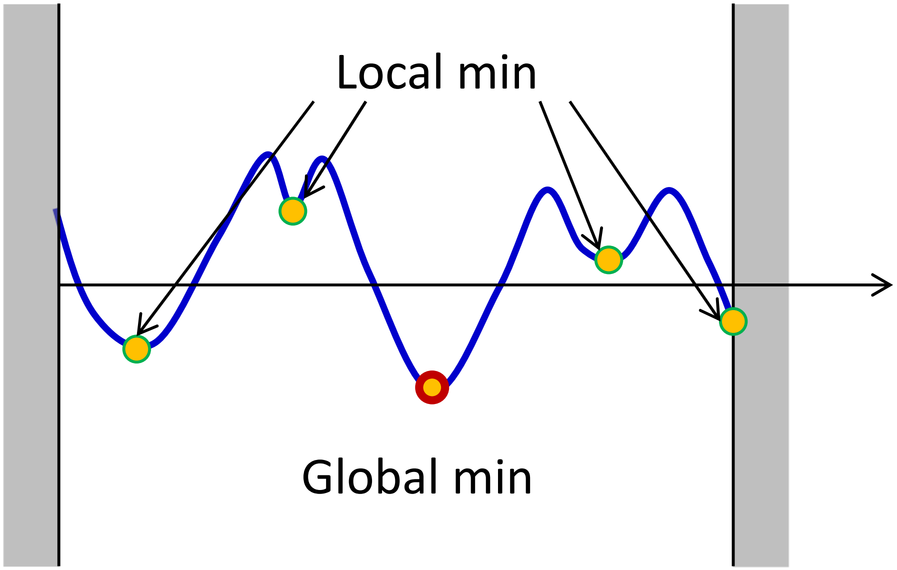
- A point **b*** is called a global minimum for *f*(**b**) if

$$f(\mathbf{b}^*) \le f(\mathbf{b}) \qquad \forall \mathbf{b} \in S$$

- No mathematical method to find the global minimum
- Weierstrass theorem: Existence of global minimum
  - If *f*(**b**) is continuous and the set *S* is closed and bounded, then there is a global minimum
- Local Optimum
  - A point **b*** is called a local minimum for *f*(**b**) if

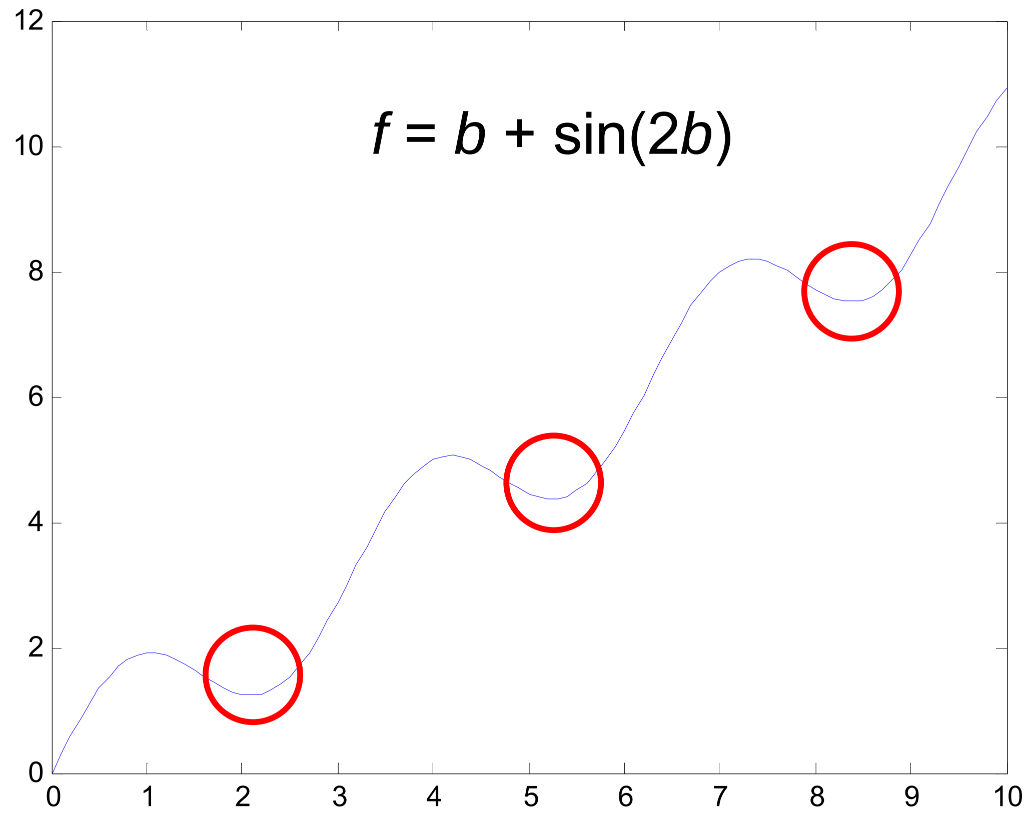
$$f(\mathbf{b}^*) \le f(\mathbf{b})$$

  - for all $\mathbf{b} \in S$ in a small neighborhood of **b***



GLOBAL OPTIMIZATION

- Normally no functional expression available
  - For a given design, we can calculate objective & constraints
- We find optima using numerical search
- We know that there is no better design in the immediate neighborhood
- But, we don't know if that is the global optimum
- We can only guarantee a local optimum

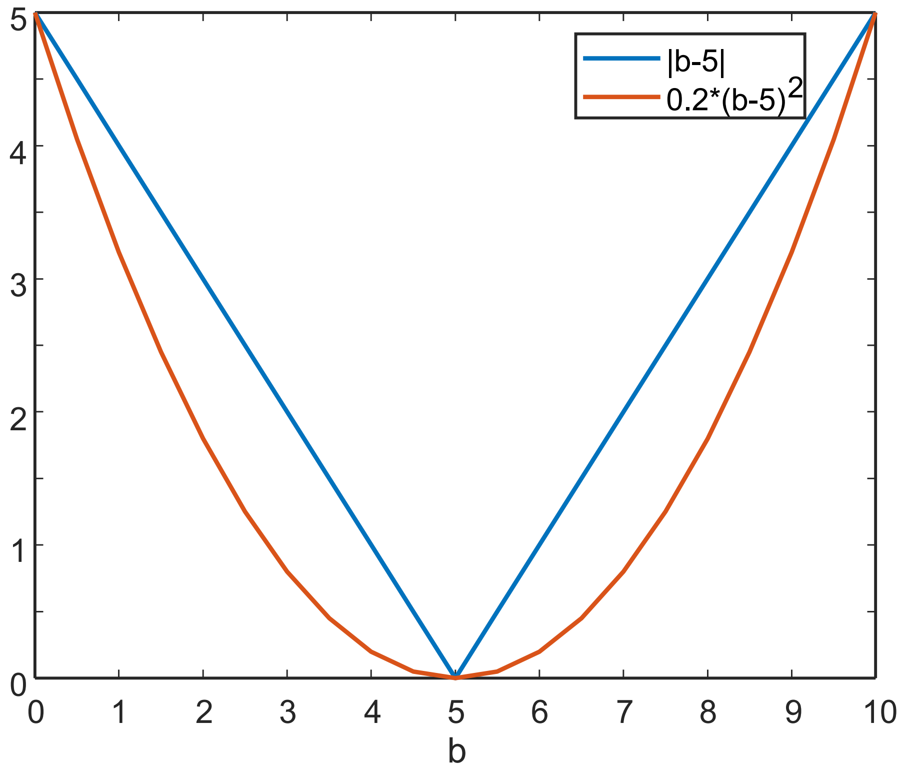


- We are accustomed to think that if $f(b)$ has a minimum then

$$\frac{\mathrm{d}f(b)}{\mathrm{d}b} = 0$$

but….

Optimality criteria only consider smooth functions



No derivative at b = 5 for a non-smooth function

1D OPTIMIZATION JARGON

- A point with zero derivative is a stationary point
- b = 5 can be
  - a minimum

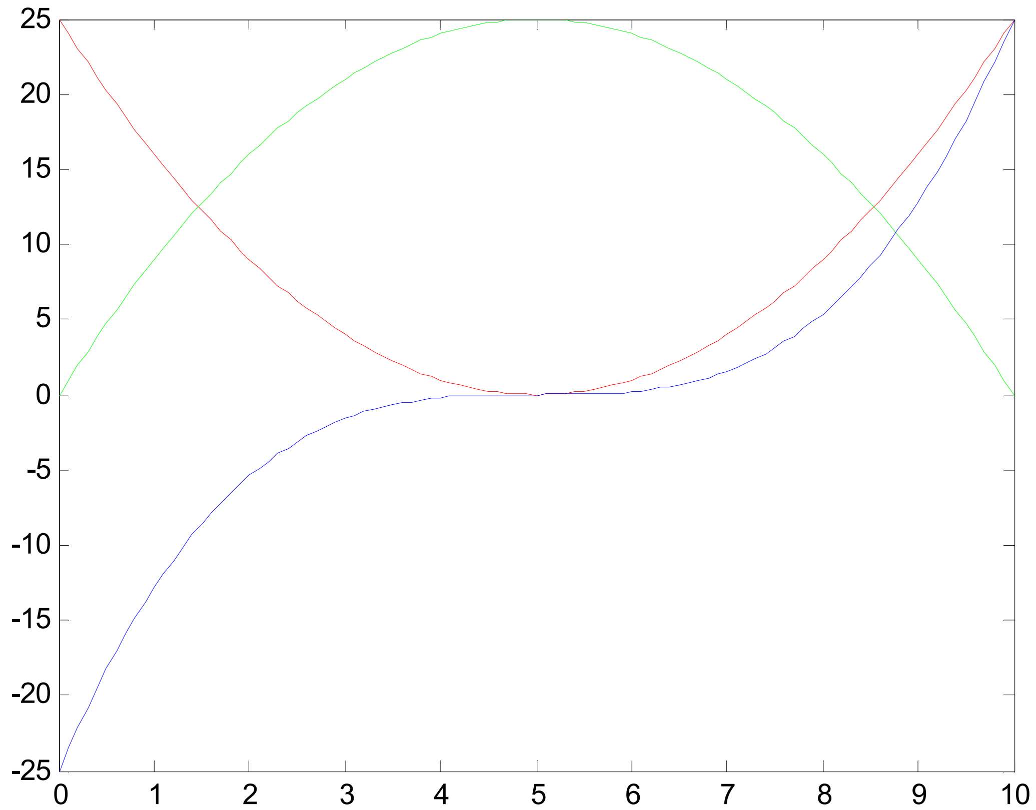
$$f = (b-5)^2$$

  - a maximum

$$f = 10b - b^2$$

  - an inflection point

$$f = 0.2(b-5)^3$$



OPTIMALITY CONDITIONS

- Unconstrained problems (one variable)

$$\underset{b}{\text{Minimize}} \quad f(b)$$

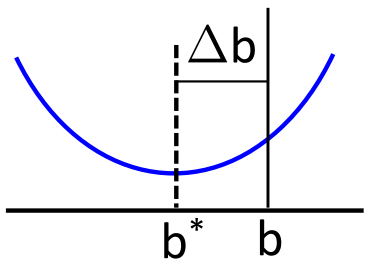
- Necessary condition for b^* to be a local min

$$\boxed{f'(b^*) = 0}$$

Kuhn-Tucker (KT) condition
1st-order necessary condition

$$f(b) = f(b^* + \Delta b) = f(b^*) + f'(b^*)\Delta b + \tfrac{1}{2} f''(b^*)\Delta b^2 + \text{H.O.T.}$$

$$\Delta f = f(b) - f(b^*) = f'(b^*)\Delta b + \tfrac{1}{2} f''(b^*)\Delta b^2 + \text{H.O.T.}$$



- For b^* to be minimum, $\Delta f \geq 0$

$$\Delta f \cong f'(b^*)\Delta b \geq 0 \ \text{ for arbitrary } \Delta b$$

$$\Rightarrow \quad f'(b^*) = 0$$

OPTIMALITY CONDITIONS *cont.*

- Now,

$$\Delta f = \tfrac{1}{2} f''(b^*)\Delta b^2 + \text{H.O.T.} \geq 0 \quad \Rightarrow \quad f''(b^*) \geq 0$$

$$\boxed{f''(b^*) \geq 0}$$

2nd-order necessary condition

- Sufficient condition

$$\boxed{f''(b^*) > 0}$$

EXERCISES

- Classify the stationary points of the following functions from the optimality conditions, then check by plotting them
  – $f(b) = 2b^3 + 3b^2$
  – $f(b) = 3b^4 + 4b^3 - 12b^2$
  – $f(b) = b^5$
  – $f(b) = b^4 + 4b^3 + 6b^2 + 4b$
- Answer true or false
  – A function can have a negative value at its maximum point
  – If a constant is added to a function, the location of its minimum point can change.
  – If the curvature of a function is negative at a stationary point, then the point is a maximum.

TAYLOR SERIES EXPANSION IN N DIMENSIONS

- Expanding $f(b_1, b_2, \ldots, b_N)$ about a candidate minimum **b***

$$f(\mathbf{b}) = f(\mathbf{b}^*) + \sum_{i=1}^{N}\left(b_i - b_i^*\right)\frac{\partial f}{\partial b_i}(\mathbf{b}^*) + \frac{1}{2}\sum_{j=1}^{N}\sum_{i=1}^{N}\left(b_i - b_i^*\right)\left(b_j - b_j^*\right)\frac{\partial^2 f}{\partial b_i \partial b_j}(\mathbf{b}^*) + \cdots$$

$$= f(\mathbf{b}^*) + \Delta\mathbf{b}^{\mathrm{T}}\nabla f(\mathbf{b}^*) + \frac{1}{2}\Delta\mathbf{b}^{\mathrm{T}}\mathbf{H}(\mathbf{b}^*)\Delta\mathbf{b} + \cdots$$

- The condition for stationarity

If $\frac{\partial f}{\partial b_i} \neq 0$ choose $\left(b_i - b_i^*\right)$ of opposite sign and other $\left(b_j - b_j^*\right) = 0$

$\Rightarrow \quad \Delta f < 0$

So must have $\nabla f = 0$

$$\nabla f = \left\{\frac{\partial f}{\partial b_1} \quad \frac{\partial f}{\partial b_2} \quad \cdots \quad \frac{\partial f}{\partial b_N}\right\}^{\mathrm{T}} = \mathbf{0}$$

CONDITIONS FOR MINIMUM

$$f(\mathbf{b}) = f(\mathbf{b}^*) + \Delta\mathbf{b}^T \nabla f(\mathbf{b}^*) + \frac{1}{2}\Delta\mathbf{b}^T \mathbf{H}(\mathbf{b}^*)\Delta\mathbf{b} + \text{H.O.T.}$$

- Sufficient condition for a minimum is that

$$\Delta\mathbf{b}^T \mathbf{H}(\mathbf{b}^*)\Delta\mathbf{b} > 0 \quad \text{for all } \Delta\mathbf{b} \neq \mathbf{0}$$

- That is, the matrix of second derivatives (Hessian) is positive definite
- Simplest way to check positive definiteness is eigenvalues: All eigenvalues need to be positive
- Necessary conditions: Hessian matrix is positive-semi definite, i.e., all eigenvalues are non-negative

OPTIMALITY CONDITION FOR UNCONSTRAINED PROBLEM

- Multi-variable case
  - KT condition

$$\nabla f(\mathbf{b}^*) = \mathbf{0}$$

  - Second-order necessary condition

$$\Delta\mathbf{b}^T\mathbf{H}(\mathbf{b}^*)\Delta\mathbf{b} \geq 0 \quad \text{for all } \Delta\mathbf{b} \in R^N$$

  - Sufficient condition

$$\Delta\mathbf{b}^T\mathbf{H}(\mathbf{b}^*)\Delta\mathbf{b} > 0 \quad \text{for all } \Delta\mathbf{b} \neq \mathbf{0}, \Delta\mathbf{b} \in R^N$$

$$\nabla f = \left\{ \frac{\partial f}{\partial b_i} \right\} \quad i = 1,\dots,N$$

Column vector

$$\mathbf{H} = \left[\frac{\partial^2 f}{\partial b_i \partial b_j} \right] \quad i,j = 1,\dots,N$$

Hessian matrix

TYPES OF STATIONARY POINTS

$$q = \Delta\mathbf{b}^T\mathbf{H}\Delta\mathbf{b}: \text{quadratic form}$$

$$q > 0 \quad \mathbf{H}: P.D \qquad q \geq 0 \quad \mathbf{H}: P.S.D$$

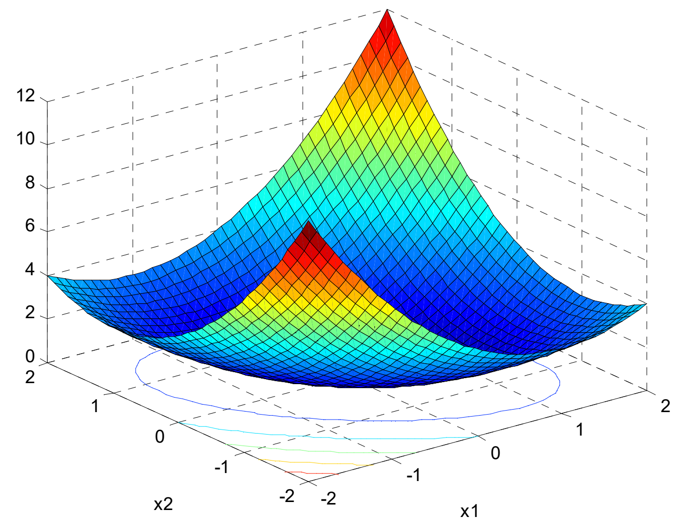
$$q < 0 \quad \mathbf{H}: N.D \qquad q \leq 0 \quad \mathbf{H}: N.S.D$$

$$\text{otherwise } \mathbf{H}: \text{indefinite}$$

- Positive definite: Minimum
- Positive semi-definite: possibly minimum
- Indefinite: Saddle point
- Negative semi-definite: possibly maximum
- Negative definite: maximum

EXAMPLE

$$f = b_1^2 + b_1 b_2 + b_2^2 \quad \nabla f = \begin{Bmatrix} 2b_1 + b_2 \\ b_1 + 2b_2 \end{Bmatrix}$$

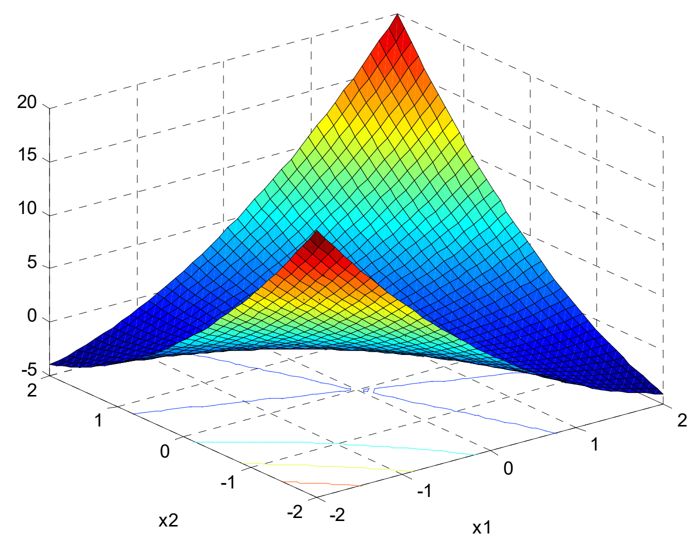


Stationary point: $b_1 = b_2 = 0$

Hessian matrix

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial b_1^2} & \frac{\partial^2 f}{\partial b_1 \partial b_2} \\ \frac{\partial^2 f}{\partial b_1 \partial b_2} & \frac{\partial^2 f}{\partial b_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Eigenvalues: $\lambda_{1,2} = 1,3 \rightarrow$ minimum



$$f = b_1^2 + 3b_1 b_2 + b_2^2$$

Eigenvalues: $\lambda_{1,2} = -1,5 \rightarrow$ saddle point

EXERCISES

- Find the stationary points of the following functions and classify them:

$$1.\, f(\mathbf{b}) = b_1^2 + 4b_1b_2 + 2b_1b_3 - 7b_2^2 - 6b_2b_3 + 5b_3^2$$

$$2.\, f(\mathbf{b}) = b_1^2 + 2b_2b_3 + b_2^2 + 4b_3^2$$

$$3.\, f(\mathbf{b}) = 40b_1 + b_1^2b_2 + \frac{b_2^2}{b_1}$$

Review of linear algebra

- Set of vectors: $\mathbf{a}^{(1)}, \cdots, \mathbf{a}^{(k)} \quad \rightarrow \quad \mathbf{a}^{(i)} \in R^n$ has n components
- Linear combination $\mathbf{b} = \sum_{i=1}^{k} x_i \mathbf{a}^{(i)}$

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \qquad \mathbf{A} = \left[\mathbf{a}^{(1)}, \cdots, \mathbf{a}^{(k)}\right]_{n \times k}$$

- Linear independence
  - From linear combination and set it to zero, $\mathbf{A}\mathbf{x} = 0$, if $\mathbf{x} = 0$ is the only solution, then columns of $\mathbf{A}$ are linear independent
  - Non-trivial solutions $(\mathbf{x} \neq 0)$ exist if $\mathbf{A}$ is rank deficient
  - If it has only $\mathbf{x} = 0$ unique solution, $rank(\mathbf{A}) = k$

Vector space

- Set of vectors must be closed under addition and scalar multiplication

$$\begin{cases} x, y \in S & \rightarrow & x + y \in S \\ \alpha \in R & \rightarrow & \alpha y \in S \end{cases}$$

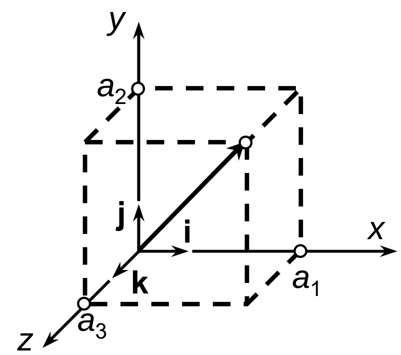
- Let $S = \{\mathbf{x} \in R^3 | x_1 = 1\}$, let $\mathbf{x}^1 = [1, a, b]^T, \mathbf{x}^2 = [1, c, d]^T$, then $\mathbf{x}^1 + \mathbf{x}^2 = [2, a + c, b + d]^T$ does not belong to S. Therefore S is not a vector space

- No. of linearly independent vectors in a set is called dimension of the vector space

- Linear independent vectors form the basis of the vector space

Orthogonality of vectors

- Orthogonal: ${\mathbf{a}^{(i)}}^T \mathbf{a}^{(j)} = 0$
- If $\left(\mathbf{a}^{(i)}, \mathbf{a}^{(j)}\right) = 0 \text{ for } \forall i \neq j$, then the set is called an orthogonal set
- If $\left\|\mathbf{a}^{(i)}\right\| = 1$, then the set is called an orthonormal set
- In $\mathbf{Ax} = \mathbf{b}$, columns of $\mathbf{A}$ form a basis for k-dim subspace



- Null space: $\mathbf{A}^T\mathbf{y} = 0$
  - Collection of all $\mathbf{y}_{(n\times 1)}$ such that $\mathbf{A}^T\mathbf{y} = 0$ is called the null space
  - No. of independent vector of null space $= n - k$

Hyperplane

- 3D plane: $ax + ay + az = d$
- n-dimension hyperplane: $a_1x_1 + a_2x_2 + \cdots a_nx_n = c$
  - Vector notation: $\mathbf{a}^T\mathbf{x} = c$ or $(\mathbf{a}, \mathbf{x}) = c$
  - Vector $\mathbf{a}$ is normal to the plane
  - Let $f(\mathbf{x}) = (\mathbf{a}, \mathbf{x}) - c$, then $\nabla f = \mathbf{a}$: gradient is normal to the surface
  - If $|\mathbf{a}| = 1$, then c is the least distance from the origin to the hyperplane
  - If $\mathbf{a}$ is not a unit vector, then the distance $= \frac{c}{\|\mathbf{a}\|}$
- If two points $\mathbf{x}, \mathbf{y}$ are on the plane, then
  - $(\mathbf{a}, \mathbf{x}) = c, (\mathbf{a}, \mathbf{y}) = c, \big(\mathbf{a}, (\mathbf{x} - \mathbf{y})\big) = 0 \;\rightarrow\; \mathbf{a}$ is normal to $\mathbf{x} - \mathbf{y}$

Hyperplane *cont.*

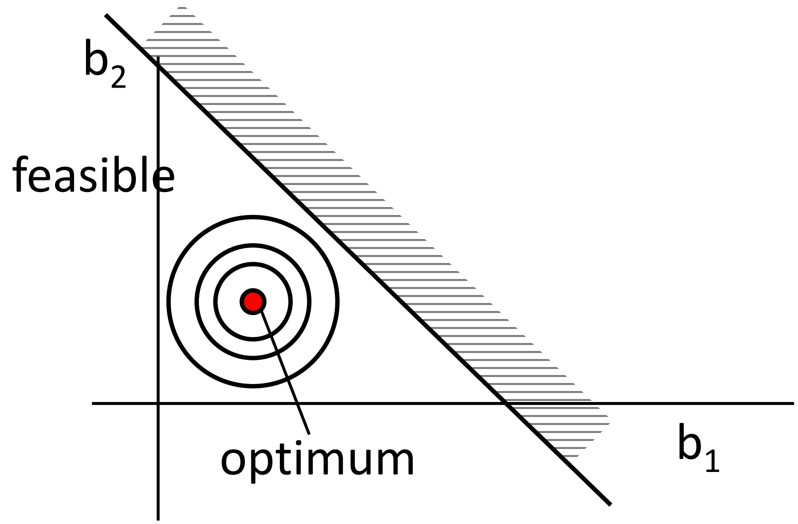
- A set of vectors satisfying $(\mathbf{a}, \mathbf{x}) = c$ is not a subspace
  - $(\mathbf{a}, \mathbf{x}) = c, (\mathbf{a}, \mathbf{y}) = c, \big(\mathbf{a}, (\mathbf{x} + \mathbf{y})\big) = 2c \quad \rightarrow$ not a subspace
- If $c = 0$ (hyperplane passes through the origin), then vectors satisfying $(\mathbf{a}, \mathbf{x}) = 0$ form a subspace
  - $(\mathbf{a}, \mathbf{x}) = 0$: vector $\mathbf{a}$ forms a basis for subspace of dimension one
  - Null space of vector $\mathbf{a}$ has dimension $n - 1$
- Reisz representation
  - Given a vector $\mathbf{x} \in R^n$, $\mathbf{x}$ can be decomposed into sum of two vectors $\mathbf{y}, \mathbf{z} \quad \rightarrow \quad \mathbf{x} = \mathbf{y} + \mathbf{z}$ when $\mathbf{y} \in F$ and $\mathbf{z} \in F^{\perp}$
  - F: subspace of m (<n) dim. $F^{\perp}$: Null space of F dim($n - m$)

CONSTRAINED PROBLEM (INACTIVE CONSTRAINT)

- Inequality constraint example

Minimize $f(\mathbf{b}) = (b_1 - 1)^2 + (b_2 - 1)^2$

subject to $g(\mathbf{b}) = b_1 + b_2 \leq 4$



- Constraint is not active → ignore

$$\frac{\partial f}{\partial b_1} = 2(b_1 - 1) = 0 \qquad b_1 = 1$$

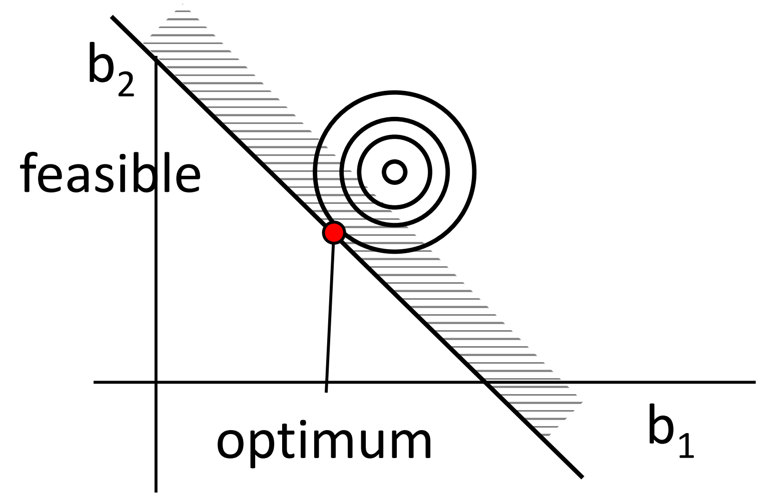
$$\frac{\partial f}{\partial b_2} = 2(b_2 - 1) = 0 \qquad b_2 = 1$$

Inactive constraints do not affect optimum

- Hessian is positive definite (sufficient) $\mathbf{H} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

CONSTRAINED PROBLEM (ACTIVE CONSTRAINT)

- Inequality constraint example



$$\text{Minimize} \quad f(\mathbf{b}) = (b_1 - 3)^2 + (b_2 - 3)^2$$
$$\text{subject to } g(\mathbf{b}) = b_1 + b_2 \le 4$$

- Constraint is active
- At (3,3), f = 0, g = 3+3−4 = 2 >0 (infeasible)
- At (2,2), f = 2, g = 0 (constraint is active)

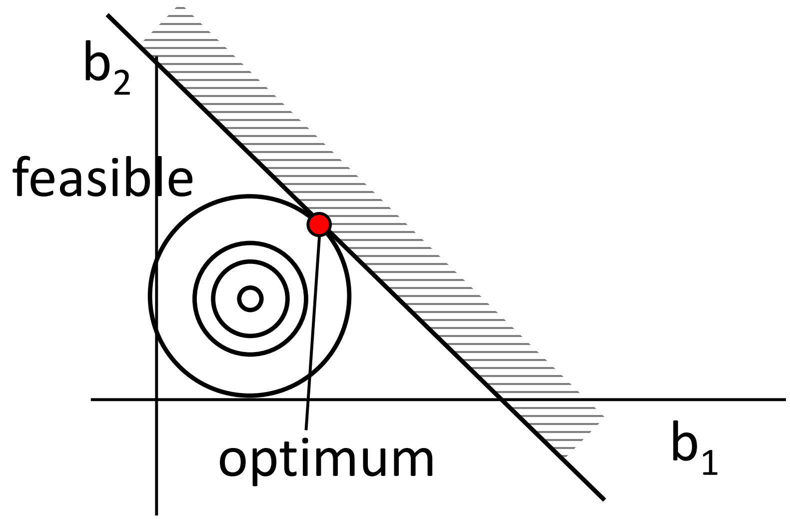
Optimum design is located on the boundary of active constraints
At optimum design, $g(\mathbf{b}) = b_1 + b_2 = 4$. Inequality becomes equality

CONSTRAINED PROBLEM (EQUALITY CONSTRAINT)

- Equality constraint example

  Minimize $f(\mathbf{b}) = (b_1 - 1)^2 + (b_2 - 1)^2$

  subject to $h(\mathbf{b}) = b_1 + b_2 = 4$



- Equality constraint is always active

- Let $b_2 = 4 - b_1$, then the original constrained problem becomes unconstrained problem with $f(b_1) = (b_1 - 1)^2 + (4 - b_1)^2$
  - Equality constraint can reduce design variables
  - If constraints are implicit, we cannot reduce design variables

OPTIMALITY CONDITION (EQUALITY CONSTRAINT)

- With an equality constraint:

$$\underset{\mathbf{b}}{\text{Minimize}} \quad f(\mathbf{b})$$
$$\text{subject to } h(\mathbf{b}) = 0$$

- Lagrange function:

$$\underset{\mathbf{b},\lambda}{\text{Minimize}} \quad \mathcal{L}(\mathbf{b},\lambda) = f(\mathbf{b}) + \lambda h(\mathbf{b})$$

Lagrange function transforms to unconstrained optimization by introducing additional variable (Lagrange multiplier)

- 1st-order necessary condition:

$$\nabla \mathcal{L}(\mathbf{b},\lambda) = \mathbf{0} \quad \Rightarrow \quad \begin{cases} \dfrac{\partial f}{\partial \mathbf{b}} + \lambda \dfrac{\partial h}{\partial \mathbf{b}} = 0 \\ h(\mathbf{b}) = 0 \end{cases}$$

- Multiple constraints: $\mathcal{L}(\mathbf{b},\lambda) = f(\mathbf{b}) + \sum_{i=1}^{M} \lambda_i h_i(\mathbf{b})$

Derivation of Lagrange multiplier (equality constraint)

* Derivation of Lagrange multiplier

- $\min f(\underline{b})$
  s.t $\underline{g}(\underline{b}) = 0$ $\quad \underline{b} \in R^n \quad \underline{g} \in R^m \quad n > m$

- partition $\underline{b} = \begin{bmatrix} \underline{y} \\ \underline{z} \end{bmatrix}$ $\quad \underline{y} \in R^m$ dependent variables
  $\underline{z} \in R^{n-m}$ independent variables.

- Due to equality constraints, conceptually $\underline{y} = \phi(\underline{z})$. but we don't know $\phi(z)$.

- $\underline{g}(\underline{y}, \underline{z}) = 0 \Rightarrow \underline{g}(\phi(\underline{z}), \underline{z}) = 0$
  $f(y, z) = f(\phi(z), z) \Rightarrow$ unconstrained problem with $z \in R^{n-m}$

- Unconstrained necessary condition.

$$\underset{1\times(n-m)}{\frac{df}{dz}(\phi(z), z)} = \underset{1\times(n-m)}{\frac{\partial f}{\partial z}} + \underset{(1\times m)\times(m\times(n-m))}{\frac{\partial f}{\partial y}\frac{\partial y}{\partial z}} = 0$$

$$\frac{dg}{dz}(\phi(z), z) = \underset{m\times(n-m)}{\frac{\partial g}{\partial z}} + \underset{(m\times m)\cdot(m\times(n-m))}{\frac{\partial g}{\partial y}\frac{\partial y}{\partial z}} = 0$$

Derivation of Lagrange multiplier (equality constraint) *cont*.

- Unconstrained necessary condition.

$$\frac{df}{dz}(\phi(z), z) = \frac{\partial f}{\partial z} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial z} = 0$$

$1\times(n-m)$ $\quad$ $1\times(n-m)$ $\quad$ $(1\times m)\times(m\times(n-m))$

$$\frac{dg}{dz}(\phi(z), z) = \frac{\partial g}{\partial z} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial z} = 0$$

$m\times(n-m)$ $\quad$ $(m\times m)\cdot(m\times(n-m))$

$$\frac{\partial g}{\partial y}\cdot\frac{\partial y}{\partial z} = -\frac{\partial g}{\partial z} \Rightarrow \frac{\partial y}{\partial z} = -\left(\frac{\partial g}{\partial y}\right)^{-1}\frac{\partial g}{\partial z}$$

$$\Rightarrow \frac{df}{dz} = \frac{\partial f}{\partial z} - \underbrace{\frac{\partial f}{\partial y}\left(\frac{\partial g}{\partial y}\right)^{-1}}_{=u^T}\frac{\partial g}{\partial z} = 0$$

$\underline{u}$: Lagrange multiplier

$$\underline{u} = -\left(\frac{\partial g}{\partial y}\right)^{-T}\left(\frac{\partial f}{\partial y}\right)^{T} \Rightarrow -\left(\frac{\partial g}{\partial y}\right)^{T}u = \left(\frac{\partial f}{\partial y}\right)^{T} \Rightarrow -u^T\frac{\partial g}{\partial y} = \frac{\partial f}{\partial y}$$

$$\Rightarrow \left\{\begin{array}{ll} \frac{\partial f}{\partial z} + u^T\frac{\partial g}{\partial z} = 0 & (n-m) \text{ Eq.} \\ \frac{\partial f}{\partial y} + u^T\frac{\partial g}{\partial y} = 0 & m \text{ Eq.} \end{array}\right\} \Rightarrow \underline{b} = \begin{bmatrix} y \\ z \end{bmatrix} \Rightarrow \boxed{\begin{bmatrix} \frac{\partial f}{\partial b} + u^T\frac{\partial g}{\partial b} = 0 \\ g = 0 \end{bmatrix}} \begin{array}{l} n \text{ Eq.} \\ m. \text{Eq.} \end{array}$$

If g_i are all independent, $\left(\frac{\partial g}{\partial y}\right)^{-1}$ exists.

unknowns
$\underline{b}$: m
$\underline{u}$: n

EXAMPLE: EQUALITY CONSTRAINT

- Ex) $\underset{\mathbf{b}}{\text{Minimize}} \quad f(\mathbf{b}) = (b_1 - 1)^2 + (b_2 - 1)^2$

  subject to $h(\mathbf{b}) = b_1 + b_2 = 4$

- Lagrange function

$$\mathcal{L}(\mathbf{b}, \lambda) = (b_1 - 1)^2 + (b_2 - 1)^2 + \lambda(b_1 + b_2 - 4)$$

- KKT conditions

$$\frac{\partial \mathcal{L}}{\partial b_1} = 2(b_1 - 1) + \lambda = 0$$

$$\frac{\partial \mathcal{L}}{\partial b_2} = 2(b_2 - 1) + \lambda = 0 \quad \Longrightarrow \quad b_1 = 2, \quad b_2 = 2, \quad \lambda = -2$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = b_1 + b_2 - 4 = 0$$

EXAMPLE: QUADRATIC FUNCTION

- Quadratic objective and constraint

$$\text{Minimize} \quad f(\mathbf{b}) = b_1^2 + 10b_2^2$$

$$\text{subject to } h(\mathbf{b}) = 100 - (b_1^2 + b_2^2) = 0$$

- Lagrangian: $\mathcal{L} = b_1^2 + 10b_2^2 + \lambda\left(100 - b_1^2 - b_2^2\right)$
- Stationarity conditions

$$\frac{\partial \mathcal{L}}{\partial b_1} = 2b_1 - 2\lambda b_1 = 0, \quad \frac{\partial \mathcal{L}}{\partial b_2} = 20b_2 - 2\lambda b_2 = 0$$

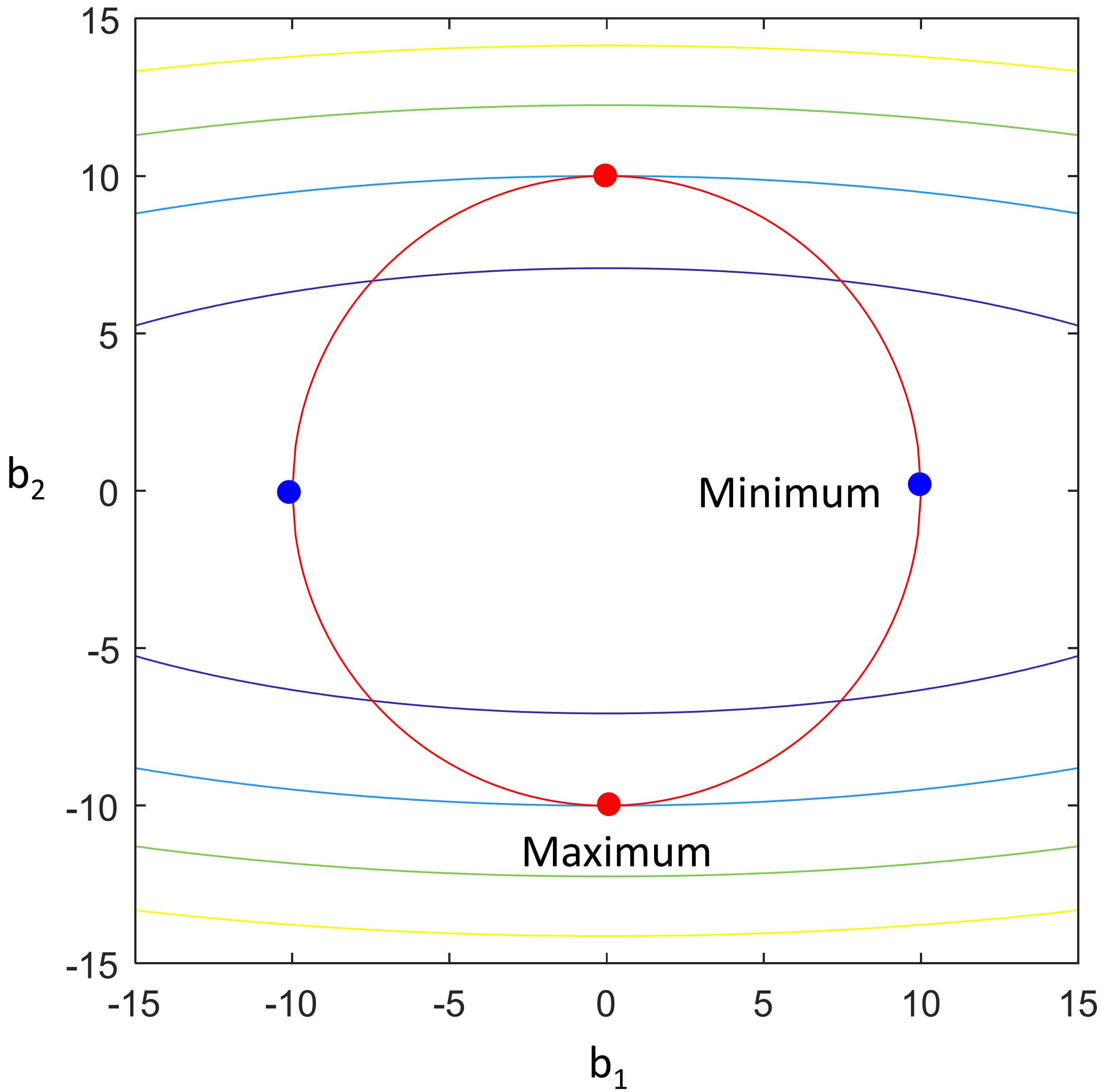
$$\frac{\partial \mathcal{L}}{\partial \lambda} = 100 - (b_1^2 + b_2^2) = 0$$

- Four stationary points

$$b_1 = 0, b_2 = \pm 10, \lambda = 10 (f = 1000, \text{ maxima})$$

$$b_1 = \pm 10, b_2 = 0, \lambda = 1 (f = 100, \text{ minima})$$

EXAMPLE: QUADRATIC FUNCTION *cont.*



EXERCISE: LAGRANGE MULTIPLIERS

- Solve the problem of minimizing the surface area of a cylinder of given volume V. The two design variables are the radius and height. The equality constraint is the volume constraint.

INEQUALITY CONSTRAINTS

- Optimization with inequality constraints

$$\text{Minimize } f(\mathbf{b}) \quad \text{subject to } g_i(\mathbf{b}) \le 0 \qquad i = 1,\dots,K$$

- Introducing a slack variable (convert to equality constraint)

$$g_i(\mathbf{b}) \le 0 \quad \Rightarrow \quad g_i(\mathbf{b}) + s_i^2 = 0 \qquad s_i^2 \ge 0 \quad \text{slack variable}$$

- Lagrange function

$$\underset{\mathbf{x},\lambda,\mathbf{s}}{\text{Minimize}} \ \mathcal{L}(\mathbf{x},\lambda,\mathbf{s}) = f(\mathbf{x}) + \sum_{i=1}^{K} \lambda_i (g_i + s_i^2)$$

$$\text{unknown}: \ \mathbf{x},\lambda,\mathbf{s} \quad (N + K + K)$$

INEQUALITY CONSTRAINTS

- KT condition

$$\nabla \mathcal{L}(\mathbf{b}, \boldsymbol{\lambda}, \mathbf{s}) = 0 \quad \Rightarrow \quad \begin{cases} \dfrac{\partial f}{\partial \mathbf{b}} + \displaystyle\sum_{i=1}^{K} \lambda_i \dfrac{\partial g_i}{\partial \mathbf{b}} = 0 & \text{N equations} \\ g_i + s_i^2 = 0 & \text{K equations} \\ 2\lambda_i s_i = 0 & \text{K equations} \end{cases}$$

Nonlinear equation

$$\lambda_i s_i = 0 \quad \Rightarrow \quad \boxed{\lambda_i g_i = 0}$$ Complementary slackness (switching cond)

- Slack variable

$\lambda_i = 0,\ g_i < 0$: inactive constraint

$\lambda_i > 0,\ g_i = 0$: active constraint

EXAMPLE: INEQUALITY CONSTRAINT

- Optimization Problem

$$\text{Minimize} \quad f(\mathbf{b}) = (b_1 - 3)^2 + (b_2 - 3)^2$$

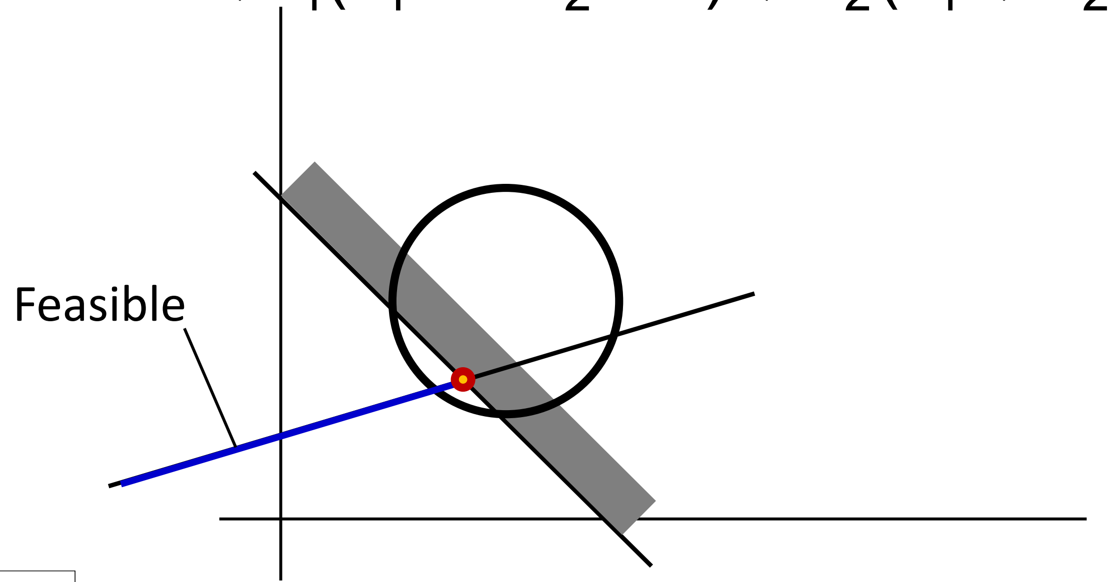
$$\text{subject to } g(\mathbf{b}) = b_1 + b_2 \le 4$$

$$h(\mathbf{b}) = b_1 - 3b_2 = 1$$

- Lagrange Function

$$\mathcal{L}(\mathbf{b}, \lambda, s) = (b_1 - 3)^2 + (b_2 - 3)^2$$

$$+\lambda_1(b_1 - 3b_2 - 1) + \lambda_2(b_1 + b_2 - 4 + s^2)$$



EXAMPLE: INEQUALITY CONSTRAINT *cont*.

$$\nabla L = \begin{cases} \dfrac{\partial L}{\partial b_1} = 2(b_1 - 3) + \lambda_1 + \lambda_2 = 0 \\ \dfrac{\partial L}{\partial b_2} = 2(b_2 - 3) - 3\lambda_1 + \lambda_2 = 0 \\ \dfrac{\partial L}{\partial \lambda_1} = b_1 - 3b_2 - 1 = 0 \\ \dfrac{\partial L}{\partial \lambda_2} = b_1 + b_2 - 4 + s^2 = 0 \end{cases}$$

Case 1) $\lambda_2 = 0$

$$2b_1 + \lambda_1 = 6$$
$$2b_2 - 3\lambda_1 = 6$$
$$b_1 - 3b_2 = 1$$
$$s^2 = 4 - b_1 - b_2$$

$\Rightarrow b_1 = 3.7,\ b_2 = 0.9,\ \lambda_1 = -1.4$

But $s^2 = -0.6$

$\therefore g_2$ is violated!

case 2) $s = 0$

$$2b_1 + \lambda_1 + \lambda_2 = 6$$
$$2b_2 - 3\lambda_1 + \lambda_2 = 6$$
$$b_1 - 3b_2 = 1$$
$$b_1 + b_2 = 4$$

$\Rightarrow b_1 = 0.75,\ b_2 = 3.25,\ \lambda_1 = -1.25$

$\lambda_2 = 0.75 > 0$.

※ $\lambda_2 \geq 0$: always positive for active inequality constraint.

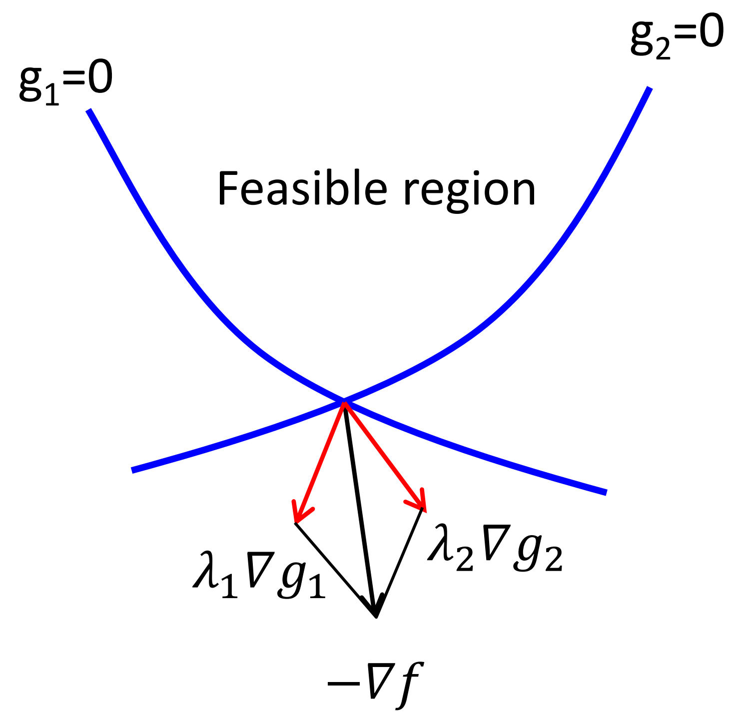
λ_1 : arbitrary in sign.

- Observation (Assume all g_i are active, i.e., $s_i = 0$)

$$\mathcal{L}(\mathbf{b},\lambda) = f(\mathbf{b}) + \sum_{i=1}^{K} \lambda_i g_i(\mathbf{b})$$

$$\nabla_{\mathbf{b}}\mathcal{L}(\mathbf{b},\lambda) = \nabla_{\mathbf{b}} f(\mathbf{b}) + \sum_{i=1}^{K} \lambda_i \nabla_{\mathbf{b}} g_i(\mathbf{b}) = 0$$

$$\nabla_{\mathbf{b}} f(\mathbf{b}) = -\sum_{i=1}^{K} \lambda_i \nabla_{\mathbf{b}} g_i(\mathbf{b})$$

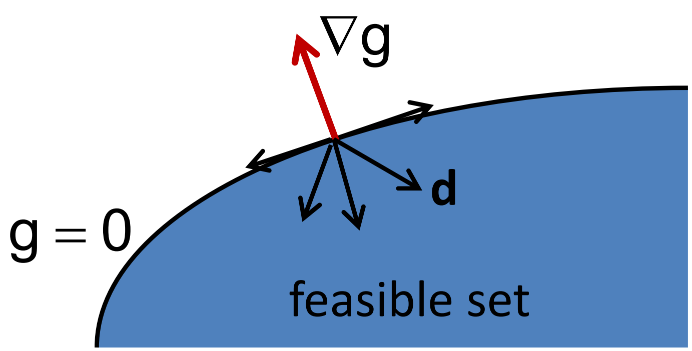


- $-\nabla_{\mathbf{b}} f(\mathbf{b})$ is a positive linear combination of $\nabla_{\mathbf{b}} g_i(\mathbf{b})$
- Only with active constraints!

- Lagrange function

$$\mathcal{L}(\mathbf{b},\boldsymbol{\mu},\boldsymbol{\lambda},\mathbf{s}) = f(\mathbf{b}) + \sum_{i=1}^{M} \mu_i h_i + \sum_{i=1}^{K} \lambda_i (g_i + s_i^2)$$

- First-order condition: $\nabla \mathcal{L}(\mathbf{b},\boldsymbol{\mu},\boldsymbol{\lambda},\mathbf{s}) = 0$
- Sufficient condition (unconstrained): $\nabla^2 f$ P.D.
- Sufficient condition (constrained): $\nabla_{\mathbf{bb}}\mathcal{L}$ P.D. for all feasible directions



Equality: $\nabla h_i \cdot \Delta \mathbf{b} = 0$

Inequality: $\nabla g_i \cdot \Delta \mathbf{b} \leq 0$

But, $\nabla g_i \cdot \Delta \mathbf{b} < 0$ direction becomes inactive

Assume that small feasible move keeps the constraint active

- Second-order necessary condition

$$q = \Delta\mathbf{b}^{\mathrm{T}}[\nabla_{\mathbf{bb}}\mathcal{L}]\Delta\mathbf{b} \geq 0 \quad \text{for all} \quad \Delta\mathbf{b} \neq \mathbf{0} \quad \text{satisfying}$$

$$\begin{cases} \nabla g_i \cdot \Delta\mathbf{b} = 0 & \text{for all active inequalities} \\ \nabla h_i \cdot \Delta\mathbf{b} = 0 & \text{for all equalities} \end{cases} \qquad (1)$$

- Second-order sufficient condition

$$q = \Delta\mathbf{b}^{\mathrm{T}}[\nabla_{\mathbf{bb}}\mathcal{L}]\Delta\mathbf{b} > 0 \quad \text{for all} \quad \Delta\mathbf{b} \neq \mathbf{0} \quad \text{satisfying (1)}$$

EFFECT OF CONSTRAINT LIMIT

- Let's work with non-standard form for the moment

$$h_i(\mathbf{b}) = a_i, \quad g_j(\mathbf{b}) \le c_j$$

a_i, c_j: constraint bounds

- Optimum design: **b*** = **b***(**a**, **c**)
- Optimum objective: $f = f(\mathbf{a}, \mathbf{c})$
- Optimum Lagrange multipliers: μ_i^* for $h_i(\mathbf{b}^*)$ and λ_i^* for $g_i(\mathbf{b}^*)$

$$\frac{\partial f}{\partial a_i} = -\mu_i^*, \quad \frac{\partial f}{\partial c_i} = -\lambda_i^*$$

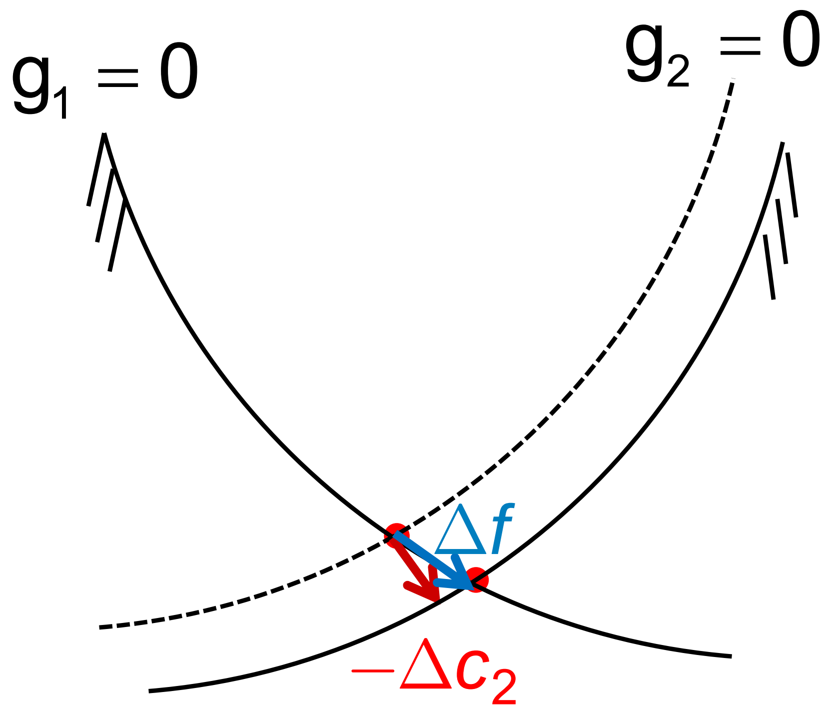
- How much will f change due to $\Delta\mathbf{a}$ and $\Delta\mathbf{c}$?

$$f(a_i + \Delta a_i, c_j + \Delta c_j) = f(a_i, c_j) + \frac{\partial f}{\partial a_i}\Delta a_i + \frac{\partial f}{\partial c_j}\Delta c_j + \text{H.O.T.}$$

$$\Delta f = -\sum_{i=1}^{M} \mu_i^* \Delta a_i - \sum_{j=1}^{K} \lambda_j^* \Delta c_j$$

EFFECT OF CONSTRAINT LIMIT *cont.*

- Change in optimal objective (Δf) due to change in constraint bound (Δc_2)



$$\Delta f = -\sum_{i=1}^{M} \mu_i^* \Delta a_i - \sum_{j=1}^{K} \lambda_j^* \Delta c_j$$

SENSITIVITY OF OPTIMUM SOLUTION TO PARAMETERS

- Assuming that objective and constraints depend on parameter p

$$\begin{aligned} &\text{Minimize} \quad f(\mathbf{b}, p) \\ &\text{subject to} \quad g_j(\mathbf{b}, p) \le 0 \end{aligned}$$

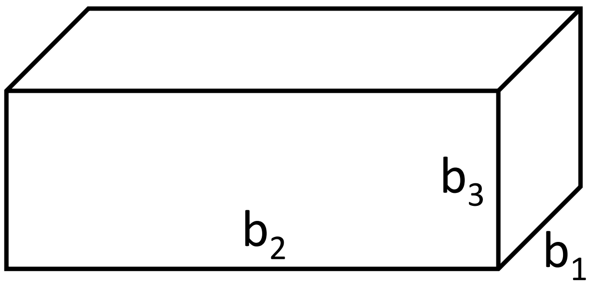
- Optimum solution: $\mathbf{b}^*(p)$
- Optimum objective: $f^*(p) = f(\mathbf{b}^*(p), p)$
- Sensitivity of $f^*(p)$ w.r.t. p

$$\frac{df^*}{dp} = \frac{\partial f}{\partial p} + \lambda^T \frac{\partial g_j}{\partial p}$$

- Lagrange multipliers are called “shadow prices” because they provide the price of imposing constraints

- Side panels (\$10/ft^2), ends & floor (\$15/ft^2), volume >= 125ft^3
- Optimization problem (cost)

$$\text{Minimize} \quad f(\mathbf{b}) = 20b_2b_3 + 30b_1b_3 + 15b_1b_2$$
$$\text{subject to } g(\mathbf{b}) = b_1b_2b_3 \geq 125$$



- Lagrange function

$$\mathcal{L}(\mathbf{b}, \lambda) = 20b_2b_3 + 30b_1b_3 + 15b_1b_2 + \lambda(125 - b_1b_2b_3)$$

- Optimum design

$$b_1^* = 4.8075, \; b_2^* = 7.2112, \; b_3^* = 3.6056$$
$$f(\mathbf{b}^*) = 1560.0, \; \lambda^* = 8.320$$

- Increase volume to 130ft^3

$$\Delta f = -\lambda^* \Delta c = -8.32 \times (-5) = 41.6$$

- Actual optimum $f^* = 1601.3, \; \Delta f = 41.3$

EXAMPLE: QUADRATIC OBJECTIVE INSIDE A CIRCLE

- Optimization problem

$$\text{Minimize} \quad f(\mathbf{b}) = b_1^2 + 10b_2^2$$
$$\text{subject to } g(\mathbf{b}) = p - (b_1^2 + b_2^2) \le 0$$

- For p = 100 we found λ = 1

$$b_1^* = \sqrt{p}, \quad b_2^* = 0, \quad f^*(p) = p, \quad \frac{df^*}{dp} = 1$$

- Which agrees with

$$\frac{df^*}{dp} = \frac{\partial f}{\partial p} + \lambda \frac{\partial g}{\partial p} = 0 + 1 \times 1$$

EXERCISE: SENSITIVITY OF OPTIMA

- For $f(b,p) = \sin b + pb,\ \ 0 \le b \le 2\pi$ find the minimum for p = 0, estimate the derivative df^*/dp, and check by solving again for p = 0.1 and comparing to finite difference derivative

- Calculate the derivative of the cylinder surface area with respect to change in volume using the Lagrange multiplier and compare to the derivative obtained by differentiating the exact solution