Lecture 6

Sufficient conditions for Finite-variable Constrained Minimization

ME260 Indian Institute of Science

Structural Optimization: Size, Shape, and Topology

G. K. Ananthasuresh

Professor, Mechanical Engineering, Indian Institute of Science, Bengaluru

suresh@iisc.ac.in

Outline of the lecture

Feasible perturbations

Second-order term in the Taylor series of an n-variable function

Sufficient conditions for constrained minimization

Bordered Hessian

What we will learn:

How to interpret feasible perturbations around a constrained local minimum

Positive definiteness of the Hessian is an overkill

How to check positive definiteness of the Hessian over the feasible perturbations

Significance of the bordered Hessian

Re-cap of KKT conditions

$$\underset{\mathbf{x}}{\text{Min}} \quad f(\mathbf{x})$$

Subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \le \mathbf{0}$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*) \boldsymbol{\lambda}^T + \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}^*) \boldsymbol{\mu}^T = \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}^*) = \mathbf{0}; \quad g_k(\mathbf{x}^*) \le 0$$

$$\mu_k \, g_k(\mathbf{x}^*) = 0; \;\; \mu_k \ge 0; \;\; k = 1, 2, \cdots, p$$

The first of KKT conditions says that the gradient of the objective function is a linear combination of the gradients of the equality and active inequality constraints.

Lagrange multipliers of inequality constraints cannot be negative; those of equality constraints can be of any sign.

Complementarity conditions (the third line) help decide if a constraint is active or inactive.

What if we maximize?

$$\underset{\mathbf{x}}{\text{Max}} \quad f(\mathbf{x})$$

Subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*)\boldsymbol{\lambda}^T + \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}^*)\boldsymbol{\mu}^T = \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}^*) = \mathbf{0}; \quad g_k(\mathbf{x}^*) \leq 0$$

$$\mu_k \, g_k(\mathbf{x}^*) = 0; \quad \mu_k \leq 0; \quad k = 1, 2, \cdots, p$$

Notice the change in the sign of the Lagrange multipliers.

Now they need to be non-positive; that is, they cannot be positive.

What if we flip the inequality sign?

$$\underset{\mathbf{x}}{\text{Min}} \quad f(\mathbf{x})$$

Subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \geq \mathbf{0}$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*) \boldsymbol{\lambda}^T + \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}^*) \boldsymbol{\mu}^T = \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}^*) = \mathbf{0}; \quad g_k(\mathbf{x}^*) \geq 0$$

$$\mu_k \, g_k(\mathbf{x}^*) = 0; \quad \mu_k \leq 0; \quad k = 1, 2, \cdots, p$$

Notice the change in the sign of the Lagrange multipliers.

Now they need to be non-positive; that is, they cannot be positive.

What if we maximize and flip the inequality sign?

$$\underset{\mathbf{x}}{\text{Max}} \; f(\mathbf{x})$$

Subject to

$$\mathbf{h}(\mathbf{x}) = \mathbf{0}$$

$$\mathbf{g}(\mathbf{x}) \geq \mathbf{0}$$

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + \nabla_{\mathbf{x}} \mathbf{h}(\mathbf{x}^*) \boldsymbol{\lambda}^T + \nabla_{\mathbf{x}} \mathbf{g}(\mathbf{x}^*) \boldsymbol{\mu}^T = \mathbf{0}$$

$$\mathbf{h}(\mathbf{x}^*) = \mathbf{0}; \quad g_k(\mathbf{x}^*) \geq 0$$

$$\mu_k \, g_k(\mathbf{x}^*) = 0; \quad \mu_k \geq 0; \quad k = 1, 2, \cdots, p$$

Notice the sign of the Lagrange multipliers.

Now they need to be non-negative again.

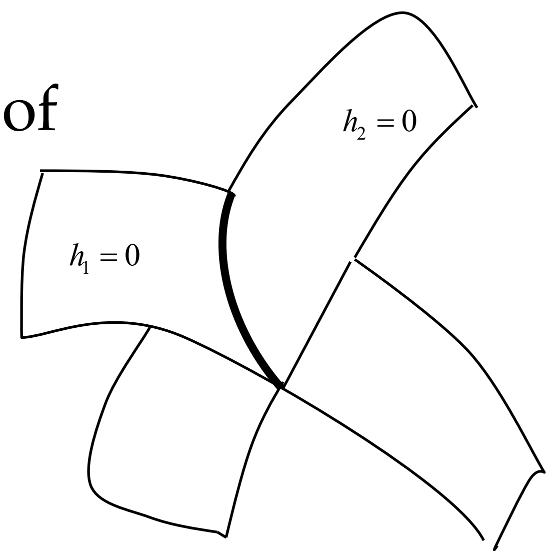
Two negatives annul each other's effect.

Feasible perturbations; constrained subspace

For sufficient conditions, we need to consider only the feasible perturbations.

Consider *m* equality constraints plus active inequality constraints such that they are linearly independent.

Together, they represent a "hyper surface" of dimension (*n*-*m*).



$$S = \left\{ \mathbf{x}^* \in \mathbb{R}^n \middle| \mathbf{h}(\mathbf{x}^*) = \mathbf{0} \right\}$$

We need to verify sufficiency by taking perturbations only in S, which is called the constrained subspace.

First order term of $f(\mathbf{x})$ in the constrained subspace

Recall from Slide 14 in Lecture 3

$$\nabla_{\mathbf{x}} f^T \Delta \mathbf{x}^* = \nabla_{\mathbf{s}} f^T(\mathbf{x}^*)\Delta \mathbf{s}^* + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*)\Delta \mathbf{d}^*$$

$$= \left\{ -\nabla_{\mathbf{s}} f^T(\mathbf{x}^*)\left[\nabla_{\mathbf{s}} \mathbf{h}^T(\mathbf{x}^*)\right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T(\mathbf{x}^*) + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*) \right\} \Delta \mathbf{d}^*$$

where $\Delta \mathbf{s}^* = -\left[\nabla_{\mathbf{s}} \mathbf{h}^T(\mathbf{x}^*)\right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T(\mathbf{x}^*)\Delta \mathbf{d}^*$

After eliminating the ***s***-variables, we can think of f as some other function z that depend only on ***d***. So, we can write in a shorthand notation:

$$\frac{\partial z}{\partial \mathbf{d}}^T = \frac{\partial f}{\partial \mathbf{s}}^T \frac{\partial \mathbf{s}}{\partial \mathbf{d}} + \frac{\partial f}{\partial \mathbf{d}}^T \quad \text{where} \quad \frac{\partial \mathbf{s}}{\partial \mathbf{d}} = -\left[\nabla_{\mathbf{s}} \mathbf{h}^T(\mathbf{x}^*)\right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T(\mathbf{x}^*)$$

Check the matrix sizes

$$\nabla_{\mathbf{x}} f^T \Delta \mathbf{x}^* = \nabla_{\mathbf{s}} f^T(\mathbf{x}^*) \Delta \mathbf{s}^* + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*) \Delta \mathbf{d}^*$$

$1\times n$ $n\times 1$ $1\times m$ $m\times 1$ $1\times(n\text{-}m)$ $(n\text{-}m)\times 1$

$$= \left\{ -\nabla_{\mathbf{s}} f^T(\mathbf{x}^*) \left[\nabla_{\mathbf{s}} \mathbf{h}^T(\mathbf{x}^*) \right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T(\mathbf{x}^*) + \nabla_{\mathbf{d}} f^T(\mathbf{x}^*) \right\} \Delta \mathbf{d}^*$$

$1\times m$ $m\times m$ $m\times(n\text{-}m)$ $1\times(n\text{-}m)$ $(n\text{-}m)\times 1$

$$\Delta \mathbf{s}^* = -\left[\nabla_{\mathbf{s}} \mathbf{h}^T(\mathbf{x}^*) \right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T(\mathbf{x}^*) \Delta \mathbf{d}^*$$

$m\times 1$ $m\times m$ $m\times(n\text{-}m)$ $(n\text{-}m)\times 1$

Second-order derivative (Hessian) of f in the constrained space

$$\frac{\partial z}{\partial \mathbf{d}}^T = \frac{\partial f}{\partial \mathbf{s}}^T \frac{\partial \mathbf{s}}{\partial \mathbf{d}} + \frac{\partial f}{\partial \mathbf{d}}^T$$

By differentiating the above first derivative, we get the second derivative.

$$\frac{d^2 z}{d\mathbf{d}^2} = \frac{d}{d\mathbf{d}}\left(\frac{\partial f}{\partial \mathbf{s}}^T \frac{d\mathbf{s}}{d\mathbf{d}}\right) + \frac{d}{d\mathbf{d}}\left(\frac{\partial f}{\partial \mathbf{d}}^T\right)$$

$$= \frac{\partial f}{\partial \mathbf{s}}^T \frac{d}{d\mathbf{d}}\left(\frac{d\bar{\mathbf{s}}}{d\bar{\mathbf{d}}}\right) + \frac{d}{d\mathbf{d}}\left(\frac{\partial f}{\partial \mathbf{s}}^T\right)\frac{d\mathbf{s}}{d\mathbf{d}} + \frac{\partial^2 f}{\partial \mathbf{d}^2} + \frac{\partial^2 f}{\partial \mathbf{d}\,\partial \mathbf{s}}\frac{d\mathbf{s}}{d\mathbf{d}}$$

$$= \frac{\partial f}{\partial \mathbf{s}}^T \frac{d^2\mathbf{s}}{d\mathbf{d}^2} + \frac{d\mathbf{s}}{d\mathbf{d}}^T \frac{\partial^2 f}{\partial \mathbf{s}\,\partial \mathbf{d}} + \frac{d\mathbf{s}}{d\mathbf{d}}^T \frac{\partial^2 f}{\partial \mathbf{s}^2}\frac{d\mathbf{s}}{d\mathbf{d}} + \frac{\partial^2 f}{\partial \mathbf{d}^2} + \frac{\partial^2 f}{\partial \mathbf{d}\,\partial \mathbf{s}}\frac{d\mathbf{s}}{d\mathbf{d}}$$

Hessian of f in the constrained space (contd.)

$$\frac{d^2 z}{d\mathbf{d}^2} = \frac{\partial f}{\partial \mathbf{s}}^T \frac{d^2\mathbf{s}}{d\mathbf{d}^2} + \frac{d\mathbf{s}}{d\mathbf{d}}^T \frac{\partial^2 f}{\partial \mathbf{s}\partial \mathbf{d}} + \frac{d\mathbf{s}}{d\mathbf{d}}^T \frac{\partial^2 f}{\partial \mathbf{s}^2}\frac{d\mathbf{s}}{d\mathbf{d}} + \frac{\partial^2 f}{\partial \mathbf{d}^2} + \frac{\partial^2 f}{\partial \mathbf{d}\partial \mathbf{s}}\frac{d\mathbf{s}}{d\mathbf{d}}$$

$$\frac{d^2 z}{d\mathbf{d}^2} = \left\{ \mathbf{I} \quad \frac{d\mathbf{s}}{d\mathbf{d}}^T \right\} \begin{bmatrix} \dfrac{\partial^2 f}{\partial \mathbf{d}^2} & \dfrac{\partial^2 f}{\partial \mathbf{d}\partial \mathbf{s}} \\ \dfrac{\partial^2 f}{\partial \mathbf{s}\partial \mathbf{d}} & \dfrac{\partial^2 f}{\partial \mathbf{s}^2} \end{bmatrix} \left\{ \begin{matrix} \mathbf{I} \\ \dfrac{\partial \mathbf{s}}{\partial \mathbf{d}} \end{matrix} \right\} + \frac{\partial f}{\partial \mathbf{s}}^T \frac{d^2\mathbf{s}}{d\mathbf{d}^2}$$

In the above expression, we know how to compute all quantities except $\dfrac{d^2\mathbf{s}}{d\mathbf{d}^2}$.

This, we will compute in the same way as we did for $\dfrac{d\mathbf{s}}{d\mathbf{d}}$, i.e., using $\mathbf{h} = \mathbf{0}$.

Hessian of the constraints in the constrained space

$\mathbf{h}=\mathbf{0}$ Requires that the second-order perturbation of the *m* constraints also be to be zero for feasibility. Therefore…

$$\frac{d^2\mathbf{h}}{d\mathbf{d}^2}=\left\{\mathbf{I} \quad \frac{d\mathbf{s}}{d\mathbf{d}}^T\right\}\begin{bmatrix}\frac{\partial^2\mathbf{h}}{\partial\mathbf{d}^2} & \frac{\partial^2\mathbf{h}}{\partial\mathbf{d}\,\partial\mathbf{s}}\\ \frac{\partial^2\mathbf{h}}{\partial\mathbf{s}\,\partial\mathbf{d}} & \frac{\partial^2\mathbf{h}}{\partial\mathbf{s}^2}\end{bmatrix}\left\{\begin{matrix}\mathbf{I}\\ \frac{\partial\mathbf{s}}{\partial\mathbf{d}}\end{matrix}\right\}+\frac{\partial\mathbf{h}}{\partial\mathbf{s}}^T\frac{d^2\mathbf{s}}{d\mathbf{d}^2}=0$$

$$\Rightarrow\frac{d^2\mathbf{s}}{d\mathbf{d}^2}=-\left[\frac{\partial\mathbf{h}}{\partial\mathbf{s}}^T\right]^{-1}\left\{\mathbf{I} \quad \frac{d\mathbf{s}}{d\mathbf{d}}^T\right\}\begin{bmatrix}\frac{\partial^2\mathbf{h}}{\partial\mathbf{d}^2} & \frac{\partial^2\mathbf{h}}{\partial\mathbf{d}\,\partial\mathbf{s}}\\ \frac{\partial^2\mathbf{h}}{\partial\mathbf{s}\,\partial\mathbf{d}} & \frac{\partial^2\mathbf{h}}{\partial\mathbf{s}^2}\end{bmatrix}\left\{\begin{matrix}\mathbf{I}\\ \frac{\partial\mathbf{s}}{\partial\mathbf{d}}\end{matrix}\right\}+$$

From Slides 10 and 11…

$$\frac{d^2 z}{d\mathbf{d}^2} = \left\{ \mathbf{I} \quad \frac{d\mathbf{s}}{d\mathbf{d}}^T \right\} \begin{bmatrix} \frac{\partial^2 f}{\partial \mathbf{d}^2} & \frac{\partial^2 f}{\partial \mathbf{d}\,\partial \mathbf{s}} \\ \frac{\partial^2 f}{\partial \mathbf{s}\,\partial \mathbf{d}} & \frac{\partial^2 f}{\partial \mathbf{s}^2} \end{bmatrix} \left\{ \begin{matrix} \mathbf{I} \\ \frac{\partial \mathbf{s}}{\partial \mathbf{d}} \end{matrix} \right\} +$$

$$+ \frac{\partial f}{\partial \mathbf{s}}^T \left(- \left[\frac{\partial \mathbf{h}}{\partial \mathbf{s}}^T \right]^{-1} \left\{ \mathbf{I} \quad \frac{d\mathbf{s}}{d\mathbf{d}}^T \right\} \begin{bmatrix} \frac{\partial^2 \mathbf{h}}{\partial \mathbf{d}^2} & \frac{\partial^2 \mathbf{h}}{\partial \mathbf{d}\,\partial \mathbf{s}} \\ \frac{\partial^2 \mathbf{h}}{\partial \mathbf{s}\,\partial \mathbf{d}} & \frac{\partial^2 \mathbf{h}}{\partial \mathbf{s}^2} \end{bmatrix} \left\{ \begin{matrix} \mathbf{I} \\ \frac{\partial \mathbf{s}}{\partial \mathbf{d}} \end{matrix} \right\} \right)$$

Recall from Slide 15 in Lecture 3 that $-\frac{\partial f}{\partial \mathbf{s}}^T \left[\frac{\partial \mathbf{h}}{\partial \mathbf{s}} \right]^{-1} = \lambda$

And now, the complete Hessian in the constrained space…

$$\frac{d^2 z}{d\mathbf{d}^2} = \begin{Bmatrix} \mathbf{I} & \dfrac{d\mathbf{s}}{d\mathbf{d}} \end{Bmatrix}^T \begin{bmatrix} \dfrac{\partial^2 L}{\partial \mathbf{d}^2} & \dfrac{\partial^2 L}{\partial \mathbf{d}\,\partial \mathbf{s}} \\ \dfrac{\partial^2 L}{\partial \mathbf{s}\,\partial \mathbf{d}} & \dfrac{\partial^2 L}{\partial \mathbf{s}^2} \end{bmatrix} \begin{Bmatrix} \mathbf{I} \\ \dfrac{\partial \mathbf{s}}{\partial \mathbf{d}} \end{Bmatrix}$$

The long expression of the last slide reduces to this because of the way we had defined the Lagrangian, L.

Where $L = f + \boldsymbol{\lambda}\mathbf{h}$

$$\Delta\mathbf{d}^{*T} \left(\frac{d^2 z}{d\mathbf{d}^2} \right) \Delta\mathbf{d}^* > 0$$

This is the sufficient condition for the constrained minimum.
Note that the perturbations are only in the independent **d** variables.

Sufficient condition for a constrained minimum

$$\Delta \mathbf{d}^{*T}\left(\frac{d^2 z}{d\mathbf{d}^2}\right)\Delta \mathbf{d}^* = \Delta \mathbf{d}^{*T}\frac{\partial^2 L}{\partial \mathbf{d}^2}\Delta \mathbf{d}^* + \Delta \mathbf{d}^{*T}\frac{\partial^2 L}{\partial \mathbf{d}\,\partial \mathbf{s}}\frac{d\mathbf{s}}{d\mathbf{d}}\Delta \mathbf{d}^* +$$

$$\Delta \mathbf{d}^{*T}\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T\frac{\partial^2 L}{\partial \mathbf{s}\,\partial \mathbf{d}}\Delta \mathbf{d}^* + \Delta \mathbf{d}^{*T}\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T\frac{\partial^2 L}{\partial \mathbf{s}^2}\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)\Delta \mathbf{d}^* > 0$$

Note: $\frac{d\mathbf{s}}{d\mathbf{d}}\Delta \mathbf{d}^* = \Delta \mathbf{s}^*$ and $\Delta \mathbf{d}^{*T}\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T = \Delta \mathbf{s}^{*T}$

Therefore, we get:

$$\Delta \mathbf{d}^{*T}\left(\frac{d^2 z}{d\mathbf{d}^2}\right)\Delta \mathbf{d}^* = \Delta \mathbf{d}^{*T}\frac{\partial^2 L}{\partial \mathbf{d}^2}\Delta \mathbf{d}^* + \Delta \mathbf{d}^{*T}\frac{\partial^2 L}{\partial \mathbf{d}\,\partial \mathbf{s}}\Delta \mathbf{s}^* +$$

$$\Delta \mathbf{s}^{*T}\frac{\partial^2 L}{\partial \mathbf{s}\,\partial \mathbf{d}}\Delta \mathbf{d}^* + \Delta \mathbf{s}^{*T}\frac{\partial^2 L}{\partial \mathbf{s}^2}\Delta \mathbf{s}^* > 0$$

Sufficient condition for a constrained minimum

$$\Delta\mathbf{d}^{*T}\frac{\partial^2 L}{\partial \mathbf{d}^2}\Delta\mathbf{d}^* + \Delta\mathbf{d}^{*T}\frac{\partial^2 L}{\partial \mathbf{d}\,\partial \mathbf{s}}\Delta\mathbf{s}^* + \Delta\mathbf{s}^{*T}\frac{\partial^2 L}{\partial \mathbf{s}\,\partial \mathbf{d}}\Delta\mathbf{d}^* + \Delta\mathbf{s}^{*T}\frac{\partial^2 L}{\partial \mathbf{s}^2}\Delta\mathbf{s}^* > 0$$

$$\Rightarrow \left\{ \Delta\mathbf{s}^* \quad \Delta\mathbf{d}^* \right\} \begin{bmatrix} \dfrac{\partial^2 L}{\partial \mathbf{d}^2} & \dfrac{\partial^2 L}{\partial \mathbf{d}\,\partial \mathbf{s}} \\ \dfrac{\partial^2 L}{\partial \mathbf{s}\,\partial \mathbf{d}} & \dfrac{\partial^2 L}{\partial \mathbf{s}^2} \end{bmatrix} \begin{Bmatrix} \Delta\mathbf{s}^* \\ \Delta\mathbf{d}^* \end{Bmatrix} > 0$$

Only feasible perturbations

$$\Rightarrow \Delta\mathbf{x}^{*T}\,\mathrm{H}(\mathbf{x}^*)\Delta\mathbf{x}^* > 0 \quad \text{with} \quad \nabla\mathbf{h}\,\Delta\mathbf{x}^* = \mathbf{0}$$

Where $\Delta\mathbf{x}^* = \begin{Bmatrix} \Delta\mathbf{s}^* \\ \Delta\mathbf{d}^* \end{Bmatrix}$

How do we check this easily?

$$\Delta\mathbf{x}^{*T}\ \mathrm{H}(\mathbf{x}^{*})\Delta\mathbf{x}^{*} > 0 \qquad \text{with} \qquad \nabla\mathbf{h}\ \Delta\mathbf{x}^{*} = \mathbf{0}$$

Note that this is a less stringent sufficient condition than requiring the positive definiteness of the Hessian, at the minimum point.

We want positive definiteness only in the subspace formed by feasible perturbations in the neighborhood of the minimum.

So, requiring positive definiteness of the Hessian is an "overkill"!

But how do we check this restricted positive definiteness? Next slide…

Bordered Hessian

$$\Delta \mathbf{x}^{*} \; \mathrm{H}(\mathbf{x}^{*})\Delta \mathbf{x}^{*} > 0 \qquad \text{with} \qquad \nabla \mathbf{h} \, \Delta \mathbf{x}^{*} = \mathbf{0}$$

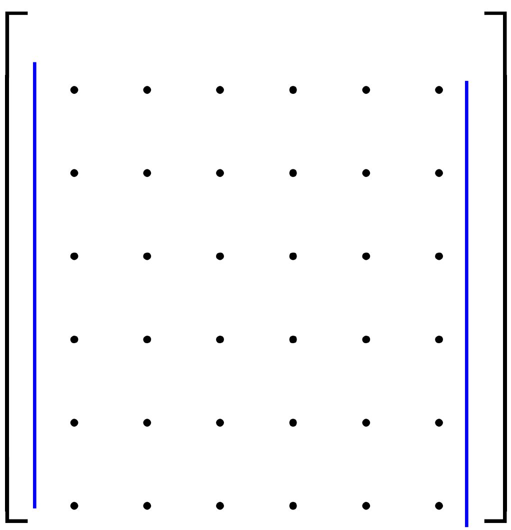
The above condition is satisfied if the last (*n*-*m*) principal minors of the **bordered Hessian, $\mathbf{H}_b$** (defined below) have the sign $(-1)^m$.

$$\mathrm{H}_b(\mathbf{x}^{*}) = \begin{bmatrix} \mathbf{0}_{m\times m} & \nabla \mathbf{h}(\mathbf{x}^{*})_{m\times n} \\ \nabla \mathbf{h}(\mathbf{x}^{*})^T_{n\times m} & H(L(\mathbf{x}^{*}))_{n\times n} \end{bmatrix}$$

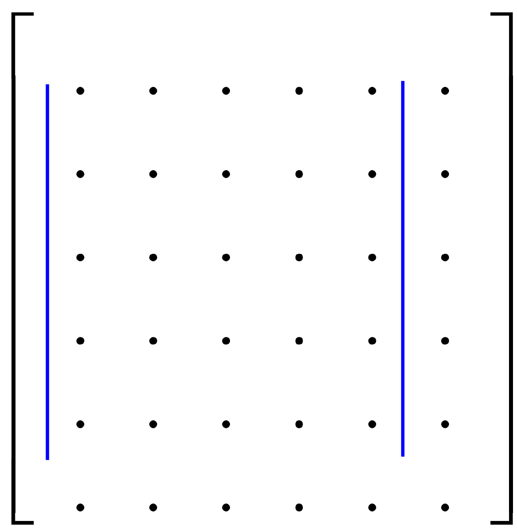
Bordered Hessian is simply Hessian of the Lagrangian bordered by the gradients of equality and active inequality constraints.

Bordered Hessian check

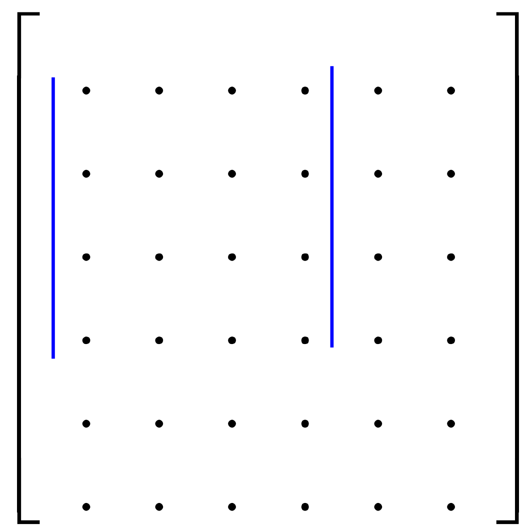
$$\mathbf{H}_b(\mathbf{x}^*) = \begin{bmatrix} \mathbf{0}_{m\times m} & \nabla \mathbf{h}(\mathbf{x}^*)_{m\times n} \\ \nabla \mathbf{h}(\mathbf{x}^*)^T_{n\times m} & H(L(\mathbf{x}^*))_{n\times n} \end{bmatrix}$$



Last principal minor



Last-but-one principal minor



Last-but-two principal minor

Example

$$\underset{x_1, x_2, x_3}{\text{Min}} \; f = x_1 + x_2^2 + x_2 x_3 + 2x_3^2$$

Subject to

$$h = 0.5(x_1^2 + x_2^2 + x_3^2) - 0.5 = 0$$

$$L = f + \lambda h = x_1 + x_2^2 + x_2 x_3 + 2x_3^2 + \lambda \left\{ 0.5(x_1^2 + x_2^2 + x_3^2) - 0.5 \right\}$$

$$\nabla L = \begin{Bmatrix} 1 + \lambda x_1 \\ 2x_2 + x_3 + \lambda x_2 \\ x_2 + 4x_3 + \lambda x_3 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

$x_1 = 1;\ x_2 = 0;\ x_3 = 0;\ \lambda = -1$ is a solution. Let us check the sufficiency.

Example (contd.)

$$\mathbf{H} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 2+\lambda & 1 \\ 0 & 1 & 4+\lambda \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix}$$

Eigenvalues of H are: -1.0000, 0.5858, and 3.4142; Not positive definite!

So, consider the Bordered Hessian:

$$\mathbf{H}_B = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & \lambda & 0 & 0 \\ 0 & 0 & 2+\lambda & 1 \\ 0 & 0 & 1 & 4+\lambda \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

$n - m = 3 - 1 = 2$; So, last two principal minors should have the sign of $(-1)^m = -1$. That is they should be negative.

Last principal minor = -2; it is fine.

Last-but-one principal minor = -1; it is also fine. So, we have a minimum.

The concept of optimization search algorithms

Optimization search algorithms work like you would walk blindfolded in a rough terrain!

They are iterative. They move from one point to another and eventually converge to a minimum at which KKT conditions are satisfied.

They need an initial guess.

Various algorithms differ in the way they choose a search direction.

Once the search direction is chosen, the algorithms needs one-variable search to decide how much to move in that direction. This is called the line search.

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{S}^{(k)}$$

Updated variable ($\mathbf{x}^{(k+1)}$); Line search parameter ($\alpha^{(k)}$); Search direction ($\mathbf{S}$); Iteration number ((k))

The end note

Sufficient conditions for
Constrained finite-variable optimization

- Recap of KKT conditions
- Feasible perturbation
  2nd order term in Taylor series expansion of an n-variable function with constraints
- Constrained subspace; Sufficient conditions for constrained minimization
  Positive definiteness of the Hessian within the constrained subspace
- **Constrained positive definiteness using bordered Hessian**
- The concept of search algorithms

Thanks