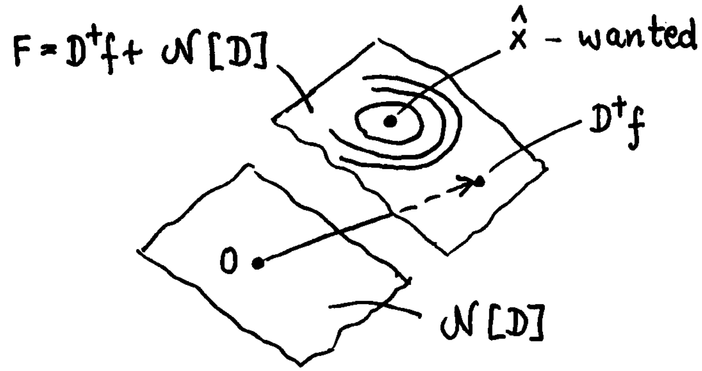
So, by Th. (11.4), if $\underline{F \neq \emptyset}$ we have this situation:

Ⓐ



Now we need an H such that:

$$\mathcal{N}[D] = \mathcal{R}[H] \leftarrow \boxed{\text{to make easy calculations with } \mathcal{N}[D]}$$

Ⓑ <u>Finding an H so that $\mathcal{N}[D] = \mathcal{R}[H]$</u>

- Apply orthogonal decomposition alg. to D:

– If: $D = P\hat{D}Q^T$ or $D = P\begin{bmatrix}\hat{D}\\ 0\end{bmatrix}Q^T$ $\Rightarrow \mathcal{N}[D] = \{0\}$ $\boxed{\text{not very likely}}$

– If: $D = P\begin{bmatrix}\hat{D}_{r\times r} & 0\\ 0 & 0\end{bmatrix}Q^T$ or $D = P\begin{bmatrix}\hat{D}_{r\times r}\\ 0\end{bmatrix}Q^T$ $\boxed{\text{more likely}}$

then: $\mathcal{N}[D] = \mathcal{R}[\underbrace{H}_{= [q_{*(r+1)} \cdots q_{*n}]}]$

So:

$$F = D^+f + \underbrace{\mathcal{N}[D]}_{\equiv \mathcal{R}[H]}$$

$$= \{D^+f + y : y \in \mathcal{R}[H]\}$$

$$= \{D^+f + H\theta : \theta \in \mathbb{R}^{n-r}\}$$

$\uparrow$ $\boxed{\text{very nice formula for a feasible set } F}$

$\Downarrow$

Hence:

$$\min_{x \in F} V(x) = \min_{x \in \{D^+f + H\theta \,:\, \theta \in \mathbb{R}^{n-r}\}} V(x)$$

constrained optimization

$$= \min_{\theta \in \mathbb{R}^{n-r}} V(D^+f + H\theta) \qquad (£)$$

unconstrained optim.
use C.G., Secant, etc.

C Solving (£) ... EASY

and then : an $\hat{x} = D^+f + H\hat{\theta}$ for any g.m. $\hat{\theta}$ of (£).

nice, easy method for solving linear equality-constrained opt. problems, by turning them into unconstrained opt. problems, by making use of the algebraic structure of F, namely using: $F = D^+f + \mathcal{R}[H]$

Nonlinear equality constrained optimization – necessary conditions for optimality

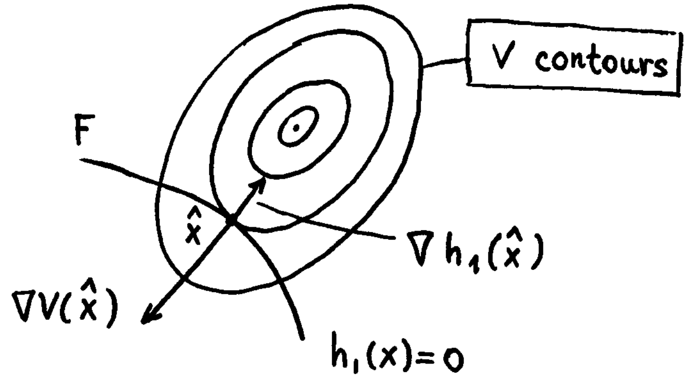
$$\min_{x \in \mathbb{R}^n} \{V(x) : h(x) = 0\} \qquad (E)$$

$h : \mathbb{R}^n \to \mathbb{R}^m$ — potentially nonlinear

e.g.

$m = 1$

$h = h_1$



$$\nabla V(\hat{x}) = -2\, \nabla h_1(\hat{x})$$

i.e. $\nabla V(\hat{x}) = -\lambda_1 \nabla h_1(\hat{x})$ for some $\lambda_1 \in \mathbb{R}$ (12.0)

necessary condition for optimality of $\hat{x}$ in (E) for the case when $m = 1$

(12.1) **Th** $\hat{x}$ is a local minimizer for (E)

Then
$$\left\{\begin{array}{l} \nabla V(\hat{x}) = -\sum_{i=1}^{m} \lambda_i \nabla h_i(\hat{x}) \\ \text{for some } \lambda_i \in \mathbb{R},\ i=1,\dots,m \\ \text{provided } \nabla h_1(\hat{x}),\dots \nabla h_m(\hat{x}) \\ \text{are linearly independent} \end{array}\right.$$

a constraint qualification, usually satisfied, so not unduly restrictive

a necessary condition for equality-constrained optimality (a general version of (12.0))

More notation

$$V_x(x) \triangleq \left[\frac{\partial V}{\partial x_1}, \dots, \frac{\partial V}{\partial x_n}\right] = \nabla V(x)^T$$

$$h_x(x) = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla h_1(x)^T \\ \vdots \\ \nabla h_m(x)^T \end{bmatrix}$$

$$\Rightarrow h_x(x)^T = [\nabla h_1(x) \dots \nabla h_m(x)]$$

Proof: Since $\hat{x}$ = a l.m. $\Rightarrow$

$V(\hat{x}+\delta x) \geq V(\hat{x})$, $\forall$ small δx such that $h(\hat{x}+\delta x) = 0$

$$h(\hat{x}+\delta x) \simeq \underbrace{h(\hat{x})}_{0} + h_x(\hat{x})\delta x \qquad (\#)$$

define: $\hat{h}_x \triangleq h_x(\hat{x})$

$\Rightarrow V(\hat{x}+\delta x) \geq V(\hat{x})$ $\forall$ small δx such that $\hat{h}_x \delta x = 0$

for subtle reasons, the accuracy of expansion (#) is not sufficient for this $\Rightarrow$ to be valid unless the $\nabla h_i(\hat{x})$ are linearly independent

but $V(\hat{x}+\delta x) \geq V(\hat{x})$ $\forall \left\{ \begin{array}{l} \text{small } \delta x \text{ s.t} \\ \hat{h}_x \delta x = 0 \end{array}\right.$

$$V(\hat{x}+\delta x) \simeq V(\hat{x}) + \nabla V(\hat{x})^T \delta x$$

for small δx

Hence :

$$\nabla V(\hat{x})^T \delta x \geqslant 0 \quad \forall \begin{cases} \text{small } \delta x \text{ such that} \\ \hat{h}_x \, \delta x = 0 \end{cases}$$

i.e., since all is now linear, then:

$$\nabla V(\hat{x})^T \delta x \geqslant 0 \quad \forall \ \delta x \text{ such that } \hat{h}_x \delta x = 0$$

$$\Rightarrow \nabla V(\hat{x})^T \delta x \geqslant 0 \quad \forall \ \delta x \in \mathcal{N}[\hat{h}_x]$$

since $(\hat{h}_x \delta x = 0) \Rightarrow (\hat{h}_x (-\delta x) = 0)$

$$\Longrightarrow \left.\begin{array}{l} \nabla V(\hat{x})^T \delta x \geqslant 0 \\ \nabla V(\hat{x})^T (-\delta x) \geqslant 0 \end{array}\right\} \forall \ \delta x \in \mathcal{N}[\hat{h}_x]$$

$$\Longrightarrow \left.\begin{array}{l} \nabla V(\hat{x})^T \delta x \geqslant 0 \\ \nabla V(\hat{x})^T \delta x \leqslant 0 \end{array}\right\} \forall \ \delta x \in \mathcal{N}[\hat{h}_x]$$

$$\Longrightarrow \nabla V(\hat{x})^T \delta x = 0 \ , \ \forall \ \delta x \in \mathcal{N}[\hat{h}_x]$$

$$\Longrightarrow \nabla V(\hat{x}) \in {}^{\perp}\mathcal{N}[\hat{h}_x] = \mathcal{R}[\hat{h}_x^T]$$

easy to prove

Hence

$$\nabla V(\hat{x}) = \hat{h}_x^T \theta \quad \text{for some } \theta \in \mathbb{R}^m$$

$$= \hat{h}_x^T (-\lambda) \quad \text{for some } \lambda \in \mathbb{R}^m$$

$$= [\nabla h_1(\hat{x}) \ldots \nabla h_m(\hat{x})](-\lambda) \quad \text{for some } \lambda \in \mathbb{R}^m$$

$$= \sum_{i=1}^{m} \nabla h_i(\hat{x}) \lambda_i$$

QED

our necessary condition for equality constrained optimality at $\hat{x}$; we also need $h(\hat{x}) = 0$, of course.

note that this condition is valid if we assume, additionally, that

h & V are C^1 functions

so the 1st-order p.ds. and expansions are OK.