Now :
- a little more of unconstrained opt.
- start of equality-constrained opt.

More notation:

Consider $h : \mathbb{R}^n \to \mathbb{R}^m$ ← $\boxed{h(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_m(x) \end{bmatrix}}$

$$h(x+\delta x) \cong h(x) + \begin{bmatrix} \sum_{i=1}^{n} \frac{\partial h_1}{\partial x_i} \delta x_i \\ \cdots \\ \sum_{i=1}^{n} \frac{\partial h_m}{\partial x_i} \delta x_i \end{bmatrix}$$

$$\equiv h(x) + \underbrace{\begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \cdots & & \cdots \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix}}_{\triangleq h_x(x) \text{ - jacobian of } h} \begin{bmatrix} \delta x_1 \\ \vdots \\ \delta x_n \end{bmatrix}$$

i.e.

$$h(x+\delta x) \cong h(x) + h_x(x)\delta x$$

validity requires all $\frac{\partial h_j}{\partial x_i}$ to exist and be continuous near x; guaranteed if h is a C^1 function meaning all 1st-order p.d.s. exist and are continuous on $\mathbb{R}^n$.

Optimization for solving : $h(x) = 0$ ← maybe nonlinear

Could solve using :

- a nonlinear vector-equation solver alg.

or by

- trying to choose x so $\|h(x)\|^2 = 0$

  i.e. solving : $\min_{x \in \mathbb{R}^n} \|h(x)\|^2$ (#)

  i.e. $\min_{x \in \mathbb{R}^n} V(x)$ with $V(x) \triangleq \|h(x)\|^2$

Formula for $\nabla V(x)$ etc.

$$V(x) = \|h(x)\|^2 = \sum_{i=1}^{n} h_i(x)^2$$

so

$$\frac{\partial V}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial}{\partial x_j} h_i^2(x) = \sum_{i=1}^{n} 2 \frac{\partial h_i}{\partial x_j}(x)$$

(#) this approach useful e.g. when fixed-point algs. cannot be used because we cannot get $\gamma < 1$

So:

$$\nabla V(x) = \begin{pmatrix} \frac{\partial V}{\partial x_1} \\ \vdots \\ \frac{\partial V}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n 2 \frac{\partial h_i}{\partial x_1} h_i \\ \vdots \\ \sum_{i=1}^n 2 \frac{\partial h_i}{\partial x_n} h_i \end{pmatrix}(x)$$

$$= 2 \underbrace{\begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_n}{\partial x_1} \\ \vdots & & \vdots \\ \frac{\partial h_1}{\partial x_n} & \cdots & \frac{\partial h_n}{\partial x_n} \end{bmatrix}}_{h_x(x)^T} \overbrace{\begin{pmatrix} h_1(x) \\ \vdots \\ h_n(x) \end{pmatrix}}^{h(x)}$$

So:

(11.1) $\nabla V(x) = 2\, h_x(x)^T h(x)$

Similarly, it can be calculated that

(11.2) $V_{xx}(x) = 2\, h_x(x)^T h_x(x) + \underbrace{2 \sum_{i=1}^n h_i(x)\, h_{i\,xx}(x)}$

$\boxed{h_i \to 0 \text{ when } V = \sum h_i^2 \text{ decreases to zero}}$

$\boxed{\text{evaluation} = \text{expensive}}$

So:

(11.2) $V_{xx}(x) \cong 2\, h_x(x)^T h_x(x)$

with the approximation getting better as $x \to$ a solution of $h(x) = 0$ since then $h_i \to 0$.

Hence: we can use the approximation

$$\bar{P}(x) = 2\, h_x(x)^T h_x(x) \quad \text{to } V_{xx}(x)$$

in a Newton-type alg.

In fact: $\bar{P}(x) \geq 0$

but is not necessarily > 0

we can, however, approximate $\bar{P}(x)$ by a positive-definite $P(x)$ by changing all zero λ_i of $\bar{P}$ to $\varepsilon > 0$

$\Rightarrow$ we can get a positive-definite approximation $P(x)$ to $V_{xx}(x)$

Once we have an approximation $P(x) > 0$ to $V_{xx}(x)$, $\forall x$ we can apply the P-alg. of § 4 to minimize $V(x)$.

A iteration j of the P-alg.:

$$s_j = -P(x_j)^{-1} \nabla V(x_j)$$

= the (so-called) <u>Gauss-Newton search direction</u>

Using this s_j at each iteration yields the Gauss-Newton Alg. for solving $h(x)=0$ by minimizing $\min_{x \in \mathbb{R}^n} \|h(x)\|^2$

Equality-constrained optimization

$$\min_{x \in F} V(x) \quad \text{where} \quad F = \{x \in \mathbb{R}^n : h(x) = 0\}$$

vector equality constraint
$h : \mathbb{R}^n \to \mathbb{R}^m$
usually $m \le n$ to help ensure that $F \neq \emptyset$

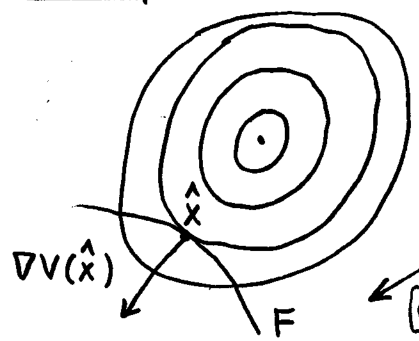
Equivalent problem formulations:

$$\min_{x \in \mathbb{R}^n} V(x) \quad \text{subject to } h(x) = 0$$

or

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

<u>Example</u>



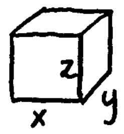
For unconstrained opt.:
$\nabla V(\hat{x}) = 0$

For constrained opt.:
$\nabla V(\hat{x}) \neq 0$

as shown here

Example

box



maximize volume

subject to : area of material $= A$ (given to us)

i.e.
$$\max_{(x,y,z)\in\mathbb{R}^3} xyz$$
subject to $2(xy+yz+xz)=A$

i.e. since $\max V$ is eqivalent to $\min(-V)$

$$\Rightarrow \min_{(x,y,z)\in\mathbb{R}^3} \{-xyz : 2(xy+yz+xz)=A\}$$

Example: $x_{k+1} = Ax_k + Bu_k \,,\; x_0 = x_0$

Minimize control effort needed to reach a desired state x_d at a time N, i.e.

$$\min_{(u_0,\dots,u_{N-1})\in\mathbb{R}^N} (u_0^2+u_1^2+\dots+u_{N-1}^2)$$

subject to :

$$x_d = A^N x_0 + A^{N-1}Bu_0 + \dots + Bu_{N-1}$$

Linear equality-constrained opt. of general V :

$$\min_{x\in\mathbb{R}^n} \{V(x) : \underbrace{D}_{\boxed{\text{a } m\times n \text{ matrix with linearly independent rows (to avoid redundant constraints)}}} x = \underbrace{f}_{\boxed{\text{general linear equality constraint}}}\}$$

Ⓘ Simple elimination method

e.g. $\min_{x\in\mathbb{R}^2} \{V(x) : ax_1+bx_2=f\}$

$\Downarrow$ — solving for x_2

$$x = \begin{pmatrix} x_1 \\ \frac{f-ax_1}{b} \end{pmatrix} \in F \,,\; \forall x_1$$

$\Downarrow$

can solve this constrained problem by:

$$\boxed{\text{unconstr. min.}} \rightarrow \min_{x_1\in\mathbb{R}} V\left(\begin{pmatrix} x_1 \\ (f-ax_1)/b \end{pmatrix}\right) \leftarrow \text{EASY}$$

this method = OK only for simple constraints.

$\Rightarrow$ a better approach is needed

II <u>The null-space method</u>

$\boxed{\text{based on a deeper understanding of } \underset{\{x \in \mathbb{R}^n : Dx = f\}}{F}}$

(11.3) **Th.** $F \neq \emptyset$ iff $DD^+ f = f$

$\boxed{\text{easy to test}}$

If $F \neq \emptyset$ then there are feasible x's so we can proceed to $\min_{x \in F} V(x)$

If $F = \emptyset$ then there are no feasible x's so we cannot $\min_{x \in F} V(x)$

$\Rightarrow$ need to re-think the constraints

<u>Proof of (11.3):</u>

$F \neq \emptyset$ iff $\exists x$ such that $f = Dx$

i.e. iff $f \in \{Dx : x \in \mathbb{R}^n\} = \mathcal{R}[D]$

i.e. iff $f \in \mathcal{R}[D]$

$$\parallel \quad \boxed{\text{orthogonal projection Th. (9.6)}}$$

$$\underset{\mathcal{R}[D]}{\underset{\cap}{\bar{f}}} + \underset{{}^{\perp}\mathcal{R}[D]}{\underset{\cap}{\tilde{f}}}$$

i.e. iff $f = \bar{f} = DD^+ f$

$\boxed{\text{Th (9.7)}}$

So $F \neq \emptyset$ iff $f \in \mathcal{R}[D]$,

i.e. iff $f = DD^+ f$

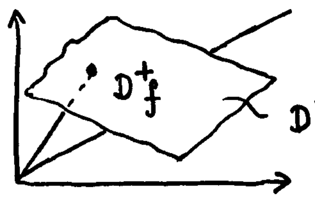
QED

More about F

(11.4) **Th** If $F \neq \emptyset$:

Then $F = D^+f + \mathcal{N}[D]$

$\triangleq \{ D^+f + y : y \in \mathcal{N}[D] \}$



Proof: Assume $F \neq \emptyset$

$\overset{\boxed{\text{Th }(11.3)}}{\Rightarrow} \quad D\underbrace{D^+f}_{x} = f \quad \Rightarrow D^+f \in F \leftarrow$ #

Proof that $D^+f + \mathcal{N}[D] \subset F$

$x \in D^+f + \mathcal{N}[D]$

$\Rightarrow x = D^+f + y$ for some $y \in \mathcal{N}[D]$

$\Rightarrow Dx = \underbrace{DD^+f}_{\overset{\#}{=} f} + \underbrace{Dy}_{=0} = f$

Hence $x \in F$. QED

Proof that $F \subset D^+f + \mathcal{N}[D]$:

$x \in F \Rightarrow Dx = f \overset{\#}{=} DD^+f$

$\Rightarrow D(\underbrace{x - D^+f}_{y \in \mathcal{N}[D]}) = 0$

$\Rightarrow x - D^+f = y$ for some $y \in \mathcal{N}[D]$

$\Rightarrow x = D^+f + y$ for some $y \in \mathcal{N}[D]$

$\Rightarrow x \in \underbrace{\{ D^+f + y : y \in \mathcal{N}[D] \}}_{\equiv D^+f + \mathcal{N}[D]}$ QED