Summary: Lagrange multipliers

(1) Minimize $f(\underline{x})$, $\underline{x} \in V$, $f: V \to \mathbb{R}$
subject to the constraints

$$c_1(\underline{x}) = 0$$
$$\vdots$$
$$c_m(\underline{x}) = 0$$

Principle:
Suppose $\underline{x}$ is a stationary point of $f(\underline{x})$ subject to these constraints.
If we perturb $\underline{x}$ in a direction allowed by the constraints, the value of $f(\underline{x})$ will not change (to first order.)

Let $\varepsilon \underline{\Delta x}$ be a proposed perturbation.
To first order, the change in the value of a function is the inner product of its derivative with the perturbation:

$$\Delta c_i \approx \varepsilon \underline{\Delta x} \cdot \nabla c_i$$

Need $\underline{\Delta x} \cdot \nabla c_i = 0 \quad \forall i$ (i.e. $\Delta x_{\parallel} = 0$, so $\Delta x = \Delta x_{\perp}$.)
for $\underline{\Delta x}$ to be allowed.

Thus:
$$\boxed{\begin{array}{l}\text{If } \Delta x \cdot \nabla c_i = 0 \ \forall i, \\ \text{Then } \Delta x \cdot \nabla f = 0\end{array}} \quad \left(\begin{array}{l}\Delta x = \Delta x_{\perp} \\ \Rightarrow \Delta x \cdot \nabla f = 0\end{array}\right)$$

But this is true iff ∇f is in the linear span of the constraints ∇c_i.
To see this, decompose ∇f and Δx into components perpendicular and parallel to the subspace spanned by the constraint gradients.

Write $\nabla f = \nabla f_{\perp} + \nabla f_{\parallel}$
$$\Delta x = \Delta x_{\perp} + \Delta x_{\parallel}$$
Then:
$$\Delta x \cdot \nabla f = \underbrace{\Delta x_{\perp}}_{\text{free}} \cdot \nabla f_{\perp} + \underbrace{\underbrace{\Delta x_{\parallel}}_{0} \cdot \nabla f_{\parallel}}_{0}$$

This will be zero ($\forall \Delta x_{\perp}$) precisely when $\nabla f_{\perp}$ is zero.
So the condition for a stationary point is:

$$\boxed{\nabla f = \lambda_1 \nabla c_1 + \dots + \lambda_m \nabla c_m} \quad \text{for some } \lambda_1, \dots, \lambda_m$$

Let $L(\underline{x}, \lambda_1, \dots, \lambda_n) = f(\underline{x}) + \lambda_1 c_1(\underline{x}) + \dots + \lambda_m c_m(\underline{x})$
L is minimized with respect to all its parameters when $\nabla L = 0$:

$$\boxed{\begin{array}{l} 0 = \nabla f + \lambda_1 \nabla c_1 + \dots + \lambda_m \nabla c_m \\ 0 = c_1 \\ \quad \vdots \\ 0 = c_m \end{array}}$$

This is precisely the set of conditions previously specified to minimize f subject to c_i.

Note: The argument above assumes that $\nabla c_i \neq 0 \ \forall i$.
At a point where $\nabla c_1 = 0$, c_1 does not constrain perturbations, so c_1 drops out of the problem.
But in this case, $\lambda_1 \nabla c_1 = 0$ for <u>any</u> λ_1.
So the boxed formulas still correctly characterize stationary points.

Note: At no point did we really assume finite dimensionality of V.
We just used existence of inner product in V and differentiability of f with respect to the inner product of V.

[Should consolidate this with development of Lagrange multipliers for the abstract problem corresponding to the problem of Lagrange multipliers for variational problems with integral constraints].

Another proof of Lagrange Multipliers

Problem:

Stationize $f(x)$ subject to

$c_1(x) = 0$

$\vdots$

$c_m(x) = 0$

Idea of solution

f and the constraints c_i are differentiable, so locally linear.

$\underline{x}_0$ is a stationary point of f subject to c_i

$\Leftrightarrow$ $\underline{x}$ is a stationary point of the local linearization of f subject to the local linearization of the constraints.

But a linear transformation has a stationary point only if it is (uniformly) constant.

Solving

Let $\underline{x}_0$ be a stationary point.

Local linearization of f:

$$f(\underline{x}) \approx \tilde{f}(\underline{x}) = f(\underline{x}_0) + (\underline{x} - \underline{x}_0) \cdot \nabla f(x_0)$$

i.e. $\Delta f \approx \Delta \tilde{f} = \Delta x \cdot \nabla f(x_0)$

Local linearization of constraints:

$$c_i(\underline{x}) \approx \tilde{c}_i(\underline{x}) = c_i(\underline{x}_0) + (x - x_0) \cdot \nabla c_i(x_0)$$

i.e. $\Delta \tilde{c}_i \approx \Delta x \cdot \nabla c_i(\underline{x}_0)$

Notation: $\Delta x \equiv x - x_0$

$\Delta f(\Delta x) \equiv f(x) - f(x_0)$

$\Delta c_i(\Delta x) \equiv c_i(x) - c_i(x_0)$

Let $\nabla c_i(x_0)^\perp$ denote the space orthogonal to $\nabla c_i(x_0)$.

(Undefined if $\nabla c_i(x_0) = 0$)

A perturbation $\Delta x (\neq 0)$ is allowed by the constraints if

$$\left.\begin{aligned} &\Delta c_i(\Delta x) = 0 \\ \Leftrightarrow\ &\Delta x \cdot \nabla c_i(x_0) = 0 \\ \Leftrightarrow\ &\Delta x \in \nabla c_i(x_0)^\perp \end{aligned}\right\} (\forall c_i)$$

i.e. $\Delta x \in \bigcap_i \nabla c_i(x_0)^\perp$

Observe that

$$\Delta \tilde{f} : \Delta x \mapsto \Delta x \cdot \nabla f(x_0)$$

$$\bigcap_i \nabla c_i(x_0)^\perp \to \mathbb{R}$$

x_0 is a stationary point

$\Leftrightarrow$ image of $\Delta \tilde{f}$ is zero

$\Leftrightarrow \nabla f(x_0) \cdot \bigcap_i \nabla c_i(x_0)^\perp = 0$

$\Leftrightarrow \nabla f(x_0) \in \left(\bigcap_i \nabla c_i(x_0)^\perp\right)^\perp$

$= \bigoplus_i \nabla c_i(x_0)$ (by corollary of lemma below)

i.e. $\nabla f(x_0)$ lies in the linear span of the $\nabla c_i(x_0)$'s.

Lemma

Let V_1 and V_2 be linear subspaces of V.

For any vector subspace $U \subset V$ let $U^\perp$ denote the orthogonal complement of U (in V).

Then $(V_1 \cap V_2)^\perp = V_1^\perp \oplus V_2^\perp$

Pf

Note: $(V_1 \cap V_2)^\perp = (V_1^\perp \oplus V_2^\perp)$ $(\forall \ldots)$

$\Leftrightarrow V_1 \cap V_2 = (V_1^\perp \oplus V_2^\perp)^\perp$ $(\forall \ldots)$

$\Leftrightarrow V_1^\perp \cap V_2^\perp = (V_1 \oplus V_2)^\perp$ $(\forall \ldots)$

But $u \in V_1^\perp \cap V_2^\perp$

$\Leftrightarrow u \in V_1^\perp$ and $u \in V_2^\perp$

$\Leftrightarrow u \cdot V_1 = 0$ and $u \cdot V_2 = 0$

$\Leftrightarrow u \cdot (V_1 \oplus V_2) = 0$

$\Leftrightarrow u \in (V_1 \oplus V_2)^\perp$.

Corollary of Lemma

$$\left(\bigcap_i V_i\right)^\perp = \bigoplus_i V_i^\perp$$