Optimization

Course Notes by Ohannes Karakashian
Transcribed and Annotated by Gregory Zitelli

Conjugate Gradient Method

Recall that for the conjugate graident method, we choose $f(u_{k+1}) = \inf_{u \in U_k + G_k} f(u)$ where G_k is the span of $\nabla f(u_0), \dots, \nabla f(u_k)$. Suppose that we know the descent direction d_k from u_k to u_{k+1}. Then the minimization problem becomes $f(u_{k+1}) = \inf_{\rho > 0} f(u_k + \rho d_k)$, which is a minimization problem in one variable. How could we find such a d_k? It turns out that if $f(v) = \frac{1}{2} v^T A v - b^T v$ is quadratic, we have the following results.

Lemma. *Let $u_{l+1} - u_l = \Delta_l = \sum_{i=0}^{l} \delta_{il} \nabla f(u_l)$ for $0 \leq l \leq k$. Then $(A\Delta_j, \Delta_i) = 0$ for all $i \neq j$. Thus, if the Δ elements are nonzero, they must be linearly independent.*

This essentially says that

$$\left(\Delta_0 \mid \Delta_1 \mid \dots \mid \Delta_k \right) = \left(\nabla f(u_0) \mid \nabla f(u_1) \mid \dots \mid \nabla f(u_k) \right) \begin{pmatrix} \delta_{00} & \delta_{01} & \dots & \delta_{0k} \\ 0 & \delta_{11} & \dots & \delta_{1k} \\ \vdots & & & \\ 0 & 0 & \dots & \delta_{kk} \end{pmatrix}$$

If the first matrix is assumed to have full rank, then the diagonal elements δ_{ii} must be nonzero up to $i = k$. Now define $d_l = \alpha \Delta_l = \alpha \sum_{i=0}^{l-1} \delta_{il} \nabla f(u_i) + \alpha \delta_{ll} \nabla f(u_l)$. Now choose $\alpha = \frac{1}{\delta_{ll}}$, so that the last term has coefficient 1. Then $d_l = \nabla f(u_l) + \sum_{i=0}^{l-1} \frac{\delta_{il}}{\delta_{ll}} \nabla f(u_i)$, so that $\lambda_{il} = \frac{\delta_{il}}{\delta_{ll}}$. Now if we can only calculate the λ terms, we will have our d_l, since the minimization problem

$$f(u_{k+1}) = \inf_{\rho \in \mathbb{R}} f(u_k - \rho d_k)$$

is solved with $\rho(u_k, d_k) = \frac{(\nabla f(u_k), d_k)}{(Ad_k, d_k)}$ since f is quadratic. It can then be shown that $\delta_{kk} = -\rho(u_k, d_k)$. Then $\lambda_{ik} = \frac{\|\nabla f(u_k)\|^2}{\|\nabla f(u_i)\|^2}$, and so we get that

$$d_k = \nabla f(u_k) + \frac{\|\nabla f(u_k)\|^2}{\|\nabla f(u_{k-1})\|^2} d_{k-1}$$

Now to describe the actual method. We start with some initial vector u_0. We define $d_0 = \nabla f(u_0)$ and $r_0 = \frac{(\nabla(u_0), d_0)}{(Ad_0, d_0)}$, and set $u_1 = u_0 - r_0 d_0$. So far, this is simply steepest decent. Assume by induction, that the algorithm has proceeded to u_k, d_{k-1}. We set $d_k = \nabla f(u_k) + \frac{\|\nabla f(u_k)\|^2}{\|\nabla f(u_{k-1})\|^2} d_{k-1}$ and $r_k = \frac{(\nabla f(u_k), d_k)}{(Ad_k, d_k)}$.

For the actual implementation, $r_0 = b = Au_0$. For $i = 1, \ldots,$ $\rho_{i-1} = r_{i-1}^T r_{i-1}$, if $i = 1$ then $d_1 = r_0$, else $\beta_{i-1} = \rho_{i-1}/\rho_{i-2}$, $d_i = r_{i-1} + \beta_{i-1} d_{i-1}$. We end this if we let $q_i = Ad_i$, $\alpha_i = \rho_{i-1}/d_i^T q_i$, $v_i = v_{i-1} + \alpha_i d_i$, and $r_i = r_{i-1} - \alpha_i q_i$, and then the convergence appears to be quick.

For the general nonlinear case, the implementation of this method is called Fletcher Reeves. Most of this theory is gone for the highly nonlinear case.

Ellipsoid Method

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and differentiable. Suppose that x^* is the minimizer of f, and x_0 is some starting vector for the method. At each iteration k, let E_k be te ellipsoid centered at x_k and suppose that $x^* \in E_k$. The hyperplane $\nabla f(x_k) \cdot (x - x_k) = 0$ seperates $\mathbb{R}^n$ into two half spaces, $S^+ = \{x : \nabla f(x_k) \cdot (x - x_k) \geq 0\}$ and $S^- = \{x : \nabla f(x_k) \cdot (x - x_k) \leq 0\}$. Since f is convex,

$$f(x) \geq f(x_k) + \nabla f(x_k) \cdot (x - x_k) \geq f(x_k)$$

for $x \in X^+$, hence we can discard S^+, so that $x^* \in S^-$. Indeed, $x^* \in E_k \cap S^-$. We let E_{k+1} be the minimum volume ellipsoid which contains $E_k \cap S^-$ and let x_{k+1} be the center of E_{k+1}.

To compute these ellipsoid, observe that every ellipsoid can be expressed as $E = \{x \in \mathbb{R}^n : (x-c)^T A^{-1}(x-c) \leq 1\}$, where c is the center and A is symmetric, positive definite. Then $x_{k+1} = x_k - \frac{1}{n+1} A_k \tilde{g}_k$, where $A_{k+1} = \frac{n^2}{n^2-1}\left(A_k - \frac{2}{n+1} A_k \tilde{g}_k \tilde{g}_k^T A_k\right)$ and $g_k = \nabla f(x_k)$, $\tilde{g}_k = \frac{g_k}{\sqrt{g_k^T A_k g_k}}$. Note that the Sherman-Morrison formula says that

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}(uV^T)A^{-1}}{1 + v^T A^{-1} u}$$

It can be shown that the volume of E_{k+1} is strictly less than $e^{-1/2n}$ times the volume of E_k, and so the volume gradually approaches zero. We obviously need to begin with an initial $A_0 = R^2 I$, with R big enough so that $x^* \in E_0$.

We would like for these centers x_k to converge to the minima. Assume that f is Lipshitz, namely $\|f(x) - f(y)\| \leq M\|x - y\|$ for some M. Fix $\epsilon > 0$ sufficiently small so that the ball $B_\epsilon = \{x : \|x - x^*\| < \frac{\epsilon}{M}\} \subseteq E_0$. Also define $S_E = \{x : f(x) \leq f(x^*) + \epsilon\} \subseteq E_0$. Suppose that $f(x_i) > f(x^*) + \epsilon$, then the Lipshitz condition means that

$$f(x) - f(x^*) \leq |f(x) - f(x^*)| \leq M\|x - x^*\| \leq \epsilon$$

which implies that $x \in S_E$, and so $B_\epsilon \subseteq S_\epsilon$. Then $\epsilon \leq e^{-k/2n^2} MR$.

For our stopping criteria, notice that $f(x^*) \geq f(x_k) + \nabla f(x_k) \cdot (x^* - x_k) \geq f(x_k) + \inf_{x \in E_k} \nabla f(x_k) \cdot (x - x_k) = f(x_k) - \sqrt{\nabla f(x_k)^T A_k \nabla f(x_k)}$. This says that after k iterations, $f(x_k) \leq f(x^*) + \sqrt{\nabla f(x_k)^T A_k \nabla f(x_k)}$. Then we can simply check the size of this square root to determine how close we are.

Projected Gradient Method

Let V be a Hilbert space, with $U \subseteq V$ nonempty, convex, and closed. Let $f : V \to \mathbb{R}$ be convex. We want $u = P(u - \rho \nabla f(u))$ for $\rho > 0$, where $P : V \to U$ is the projection onto U. Let $g(u) = P(u - \rho \nabla f(u))$, so that we are in fact looking for the fixed points of g. We impliment the fixed point algorithm $u_{k+1} = g(u_k)$. We say that the projected gradient method is simply $u_{k+1} = P(u_k - \rho_k \nabla f(u_k))$ for some choice of $\rho_k > 0$.

Theorem. *Let $U \subseteq V$ be a nonempty, closed, convex subset of a Hilbert space, with $f : V \to \mathbb{R}$ differentiable and with constants $\alpha, M > 0$ such that $(\nabla f(v) - \nabla f(u), v - u) \geq \alpha\|v - u\|^2$ and $\|\nabla f(u) - \nabla f(v)\| \leq M\|v - u\|$ for all $u, v \in V$. Suppose also that there exist a, b such that $0 < a \leq \rho_k \leq b < \frac{2\alpha}{M^2}$, then the projection gradient method with ρ_k steps converges and there exists $\beta < 1$ depending on α, M, a, b such that $\|u_{k+1} - u\| \leq \beta\|u_k - u\|$.*

Penalty Method

Theorem. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be continuous, coercive, and strictly convex. Let $U \subseteq \mathbb{R}^n$ be nonempty, closed, and convex. We let $\Psi \geq 0$ be a penalty function on $\mathbb{R}^n$, such that $\Psi(u) = 0$ if and only if $u \in U$. Then for any $\epsilon > 0$ there exists a unique u_ϵ such that $f_\epsilon(u_\epsilon) = \inf_{v \in \mathbb{R}^n} f(v)$, where $f_\epsilon(v) = f(v) + \frac{1}{\epsilon}\Psi(v)$. Furthermore, $\lim_{\epsilon \to 0} u_\epsilon$ is the minimizer of f over U.*

Proof. The minimization problem for f has a unique solutions since f is coercive. Furthermore, for each $\epsilon > 0$ the function f_ϵ is also coercive and strictly convex, so it has a unique solution u_ϵ. Then

$$f(u_\epsilon) \leq f(u_\epsilon) + \frac{1}{\epsilon}\Psi(u_\epsilon) \leq f_\epsilon(u_\epsilon) \leq f_\epsilon(u) \leq f(u) + \frac{1}{\epsilon}\Psi(u) \leq f(u)$$

where $u \in U$. Then u_ϵ is bounded since f is coercive, and so there is a subsequence of u_ϵ which converges in $\mathbb{R}^n$. The continuity of Ψ implies that this limit lies in U. ■

Kuhn-Tucker Theorem

Lemma (Farkas Lemma)**.** *Suppose that V is a vector space, and $\{w \in V : (a_i, w) \geq 0, i \in I\} \subseteq \{w \in V : (b, w) \geq 0\}$ for some a_i, then there exists λ_i so that $b = \sum_{i \in I} \lambda_i a_i$.*

Suppose that U represents a constraint set that is not necessarily convex, thought of as $U = \{v \in V : \varphi_i(v) \leq 0, i = 1, \ldots, m\}$. For $u \in U$, the cone $C(u)$ of feasible directions is the union of $\{0\}$ and the set of vectors $w \in V$ for which there exists a sequence $\{u_k\}_{k \geq 0}$ such that $u_k \in U$, $u_k \neq u$, $\lim_{k \to \infty} u_k = u$, and $\lim_{k \to \infty} \frac{u - u_k}{\|u - u_k\|} = \frac{w}{\|w\|}$. $C(u)$ is a cone with vertex 0. We also define $I(u) = \{i = 1, \ldots, m : \varphi_i(u) = 0\}$.

Theorem. *Let $U \subseteq V$ be nonempty, $(V, (\cdot, \cdot))$ a Hilbert space.*

(i) At every $u \in U$, $C(u)$ is closed.

(ii) Let $f : \Omega \to \mathbb{R}$ with $U \subseteq \Omega$ open. If f has a local minimum at u with respect to U and f is differentiable at u, then $f'(u)(v - u) \geq 0$ for all $v \in u + C(u)$.

Definition. For $u \in U$, let $C^*(u) = \{w \in V : \varphi_i'(u)w \leq 0\}$.

Definition. We say that the constraints $\{\varphi_i\}_{i=1}^m$ are qualified at the point $u \in U$ if either all φ_i for $i \in I(u)$ are affine, or there exists some $\tilde{w} \in V$ such that for $i \in I(u)$, $\varphi_i'(u)\tilde{w} \leq 0$ and $\varphi_i'(u)\tilde{w} < 0$ if φ_i is not affine.

Theorem. *Let $u \in U$, φ_i differentiable for $i \in I(u)$.*

(i) $C(u) \subseteq C^*(u)$

(ii) If the constraints are qualified at u, and if φ_i for $i \in I(u)$ are continuous at u, then $C(u) = C^*(u)$.

Theorem (Kuhn-Tucker)**.** *Let V be a Hilbert space, $U \subseteq V$ given by $U = \{v \in V : \varphi_i(v) \leq 0\}$. Let φ_i for $i \in I(u)$ be differentiable at u, with the rest of the φ_i continuous. Suppose also that a function $f : V \to \mathbb{R}$ is differentiable at u. If f has a local minimum at u with respect to U, then there exist numbers $\lambda_i(u)$ for $i \in I(u)$ such that*

$$f'(u) + \sum_{i \in I(u)} \lambda_i(u)\varphi_i'(u) = 0$$

and furthermore that $\lambda_i(u) \geq 0$.

Sufficient Conditions for a Local Minimum

Definition. Let $\Omega \subseteq V$ be convex, and φ_i convex. We say that the constraints are qualified at $u \in U$ if either the φ_i are all affine, or if there exists $\tilde{v} \in \Omega$ such that $\varphi_i(\tilde{v}) \leq 0$ and $\varphi_i(\tilde{v}) < 0$ if φ_i is not affine.

Theorem. *Let $f : \Omega \to \mathbb{R}$ with Ω convex, $U = \{v \in \Omega : \varphi_i(v) \leq 0\} \subseteq \Omega$ with φ_i convex, and $u \in U$ with f and φ_i differentiable at u.*

(i) If f has a local minimum at u with respect to U and if the constraints are qualified, then there exists $\lambda_i(u)$ which satisfy the Kuhn-Tucker condition.

(ii) Conversely, if $f : U \to \mathbb{R}$ is convex and there exist numbers λ_i such that the Kuhn-Tucker conditions are satisfied, then f has a global minimum at u with respect to U.

In practice, the Kuhn-Tucker conditions are very hard to implement. Consequently, we rely on duality to solve this infimum problem. We begin by introducing the Lagrangian $L(v, \mu) = f(v) + \sum_{i=1}^{m} \mu_i \varphi_i(v)$. Given some $\mu \in \mathbb{R}^m_+$, we define $u_\mu \in V$ as the minimizer of $L(u_\mu, \mu)$, called $G(\mu)$. The dual problem to our original problem is finding the maximum of G over $\mathbb{R}^m_+$.

Definition. Let V and M be two sets, and let $L : V \times M \to \mathbb{R}$ be a function. The point $(u, \lambda) \in V \times M$ is called a saddle point of L if

$$\sup_{\mu \in M} L(u, \mu) = L(u, \lambda) = \inf_{u \in V} L(v, \lambda)$$

Theorem. *If (u, λ) is a saddle point of $L : V \times M \to \mathbb{R}$, then*

$$\sup_{\mu \in M} \inf_{v \in V} L(v, \mu) = L(u, \lambda) = \inf_{v \in V} \sup_{\mu \in M} L(v, \mu)$$

Theorem.

(i) If $(u, \lambda) \in V \times \mathbb{R}^m_+$ is a saddle point of the Lagrangian L, then u is a solution of the minimization problem over f.

(ii) Sppose f and φ_i are convex, and that f is differentiable at $u \in U$ with qualified constraints. Then if u is the single minimizer, there exists $\lambda \in \mathbb{R}^m_+$ such that (u, λ) is a saddle point of L.

Theorem. *Suppose that φ_i are continuous and that for every $\mu \in \mathbb{R}^m_+$, the problem of finding u_μ to minimize $\inf_{v \in V} L(v, \mu)$ has a unique solution which depends continuously on μ. Then if $\lambda \in \mathbb{R}^m_+$ is any solution to the supremum over G, the corresponding u_λ is a solution the the original minimization problem over f.*

Example. Consider $f(v) = 2v_1^2 + 2v_1v_2 + 3v_2^2$, with constraint $U = \{v \in \mathbb{R}^2 : v_1 + v_2 \geq 2\}$. The Lagrangian is $L(v, mu) = 2v_1^2 + 2v_1v_2 + 3v_2^2 + \mu(2 - v_1 - v_2)$. For fixed μ, $L(v, \mu)$ is elliptic. Then the second derivative of L is

$$L''(v, \mu) = \begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix}$$

which is symmetric positive-definite. Therefore, $L(v, \mu)$ has a unique minimizer u_μ over $\mathbb{R}^2$. Solving for the minimum gives $u_1 = \frac{\mu}{5}$ and $u_2 = \frac{\mu}{1}0$. Letting $G(\mu) = L(u_\mu, \mu) = -\frac{3}{20}\mu^2 + 2\mu$, which is maxmized for $\mu = \frac{20}{3}$. Then the minimum for the original problem is $u_1 = \frac{4}{2}$ and $u_2 = \frac{2}{3}$.

Example. As another example, let $f(v) = \frac{1}{2}(Av, v) - (b, v)$, with A symmetric positive-definite and $b \in \mathbb{R}^n$. Suppose the constraint set is $\{v \in \mathbb{R}^n : Cv \leq d\}$ for C an $m \times n$ matrix and $d \in \mathbb{R}^m$. Then $\varphi_i(v) = -d_i + \sum_{j=1}^n c_{ij}v_j$ for $i = 1, \ldots, m$, and $L(v, \mu) = \frac{1}{2}(Av, v) - (b, v) + \sum_{i=1}^m \mu_i \left(-d_i + \sum_{j=1}^n c_{ij}v_j\right) = \frac{1}{2}(Av, v) - (b - C^T\mu, v) - (\mu, d)$. This is indeed elliptic, so $L(v, \mu)$ has a unique minimizer u_μ. We get that $u_\mu = A^{-1}(b - C^T\mu)$. Rather than maximizing G, we will actually attempt to minimize $-G(\mu) = \frac{1}{2}(CA^{-1}C^T\mu, \mu) - (CA^{-1}b - d, \mu) + \frac{1}{2}(A^{-1}b, b)$. Note that $CA^{-1}C^T$ is nonnegative definite, and will be positive definite only if $\ker C^T = \{0\}$. If $CA^{-1}C^T$ is positive definite then $-G$ is elliptic and therefore has a unique minimizer in $\mathbb{R}^m_+$.

Uzawa's Method

The basis of Uzawa's method is to apply the projected gradient method to the problem of $\mu \mapsto G(\mu) = \inf_{v \in V} L(v, \mu)$. Given $\lambda^0 \in \mathbb{R}^m_+$, a sequence of pairs $(u^k, \lambda^{k+1}) \in V \times \mathbb{R}^m_+$ is deined by first calculating u^k as

$$f(u^k) + \sum_{i=1}^m \lambda_i^k \varphi(u^k) = \inf_{v \in V} \left\{ f(v) + \sum_{i=1}^m \lambda_i^k \varphi_i(v) \right\}$$

and then calculating λ^{k+1} as

$$\lambda_i^{k+1} = \max\left\{\lambda_i^k + \rho\varphi_i(u^k), 0\right\}$$

Note that $\lambda^{k+1} = P_+^m\left[\lambda^k + \rho\nabla G(\lambda^k)\right]$.

Theorem (Convergence of Uzawa's Method)**.** *Suppose $V = \mathbb{R}^n$, and f is ellipc with $\alpha > 0$ its ellipticity constant. Suppose $U \subseteq V$ is of the form $U = \{v \in \mathbb{R}^n : Cv \leq d\}$, for $C \in \mathbb{R}^{m\times n}$ and $d \in \mathbb{R}^m$. Assume that U is nonempty, and let $\|C\|$ be the normal operator norm. If $0 < \rho < \frac{2\alpha}{\|c\|^2}$, then the sequence $\{u^k\}$ will converge to the unique solution of the original problem. If additionally the rank of C is m, then the sequence λ^k is also convergent to a unique solution λ of the dual problem.*

Proof. If f is elliptic (coercive), then the fact that U is closed implies that the problem has a unique solution u. Also, $L(\cdot, \mu)$ is strictly convex and coercive, so the minimization problem is Uzawa's method has unique solutions u^k. Then

$$L(, v\mu) = f(v) + (C^T\mu, v) - (\mu, d)$$

By the theorem, there exists $\lambda \in \mathbb{R}^m_+$ such that (u, λ) is a saddle point of L a minimizer of $L(u, \lambda) = \inf_{v\in V} L(v, \lambda)$. Then $\nabla f(u) + C^T\lambda = 0$. Also, $L(u, lambda) = \sup_{\mu\in\mathbb{R}^m_+} L(u, \mu)$. Since U is convex, the angle condition holds. The supremum implies that $-(\rho\varphi(u), \mu - \lambda) \leq 0$ for any $\rho > 0$, which is equivalent to

$$(\lambda - (\lambda + \rho\varphi(u)), \mu - \lambda) \geq 0$$

for all $\mu \in \mathbb{R}^m_+$, or similarly that $\lambda = P_+(\lambda+\rho\varphi(u))$. Uzawa's method sets $\nabla f(u^k)+ C^T\lambda^k = 0$, and $\lambda^{k+1} = P_+(\lambda^k + \rho\varphi(u^k))$. Subtracting these, we get that

$$\nabla f(u^k) - \nabla f(u) + C^T(\lambda^k - \lambda) = 0$$

$$\|\lambda^{k+1} - \lambda\| \leq \|(\lambda^k + \rho\varphi(u^k)) - (\lambda + \rho\varphi(u))\|$$

We can then write that

$$\begin{aligned}\|\lambda^{k+1} - \lambda\|^2 &\leq \|\lambda^k - \lambda\|^2 - 2\rho(C^T(\lambda^k - \lambda), u^k - u) + \rho^2\|C(u^k - u)\|^2 \\ &= \|\lambda^k - \lambda\|^2 - 2\rho(\nabla f(u^k) - \nabla f(u), u^k - u) + \rho^2\|C(u^k - u)\|^2 \\ &\leq \|\lambda^k - \lambda\|^2 - \rho\left(2\alpha - \rho\|C\|^2\right)\|u^k - u\|^2\end{aligned}$$

Since $0 < \rho < \frac{2\alpha}{\|C\|^2}$, it follows that $0 \leq \|\lambda^{k+1} - \lambda\| \leq \|\lambda^k - \lambda\|$. Then the sequence of norms $\|\lambda^k - \lambda\|$ is non-increasing and bounded from below, so it must converge. Then

$$\lim_{k\to\infty} \|\lambda^{k+1} - \lambda\| - \|\lambda^k - \lambda\| = 0$$

But this, together with what we saw with our long inequality, shows that

$$\rho\left(2\alpha - \rho\|C\|^2\right)\|u^k - u\|^2 \leq \|\lambda^k - \lambda\|^2 - \|\lambda^{k+1} - \lambda\|^2 \to 0$$

Then $u^k \to u$.

Next, the fact that $\|\lambda^k - \lambda\| \to 0$ implies that λ^k is a bounded sequence. Then there is a convergence subsequence $\lambda^{k'} \to \lambda' \in \mathbb{R}^m_+$. By continuity, $\nabla f(u) + C^T\lambda' = \lim_{k'\to\infty}\left(\nabla f(u^{k'}) + C^T\lambda^{k'}\right) = 0$, so that $C^T(\lambda' - \lambda) = 0$. If the range of C is m, then the kernel of C^T is zero, so that $\lambda' = \lambda$.

In this case, we wish to show that the entire sequence λ^k converges to λ. This proof follows as in the proof for the penalty function convergence. ■

Linear Programming

We now look at a common problem in linear programming, minimizing a linear function f subject to affine equality and inequality constraints. Specifically, we wish to find $u \in U$ such that $f(u) = \inf_{v\in U} f(v)$, where U is described as those $v \in \mathbb{R}^n$ such that $\sum_{j=1}^n c_{ij}v_j \leq d_i$ for $i = 1, \ldots, p$, and $\sum_{j=1}^n c_{ij}v_j = d_i$ for $i = p+1, \ldots, m$. It is even possible to transform some nonlinear problems into this form.

These kinds of problems usually arrive in three equivalent forms (note that the values of n, m will certainly be different if a problem is transformed from one type or problem to the other):

(i) Find $u \in U = \{v \in \mathbb{R}^n : Cv \leq d\}$, where $C \in \mathbb{R}^{m\times n}$, $d \in \mathbb{R}^m$, such that $f(u) = \inf_{v\in U} f(v)$ where $f(v) = (a, v)$ for $a \in \mathbb{R}^n$.
(ii) Find $u \in U = \{v \in \mathbb{R}^n_+ : Cv \leq d\}$, where $C \in \mathbb{R}^{m\times n}$, $d \in \mathbb{R}^m$, such that $f(u) = \inf_{v\in U} f(v)$ where $f(v) = (a, v)$ for $a \in \mathbb{R}^n$.
(iii) Find $u \in U = \{v \in \mathbb{R}^n_+ : Cv = d\}$, where $C \in \mathbb{R}^{m\times n}$, $d \in \mathbb{R}^m$, such that $f(u) = \inf_{v\in U} f(v)$ where $f(v) = (a, v)$ for $a \in \mathbb{R}^n$.

To transform a problem from type (i) to type (ii), we express $v \in \mathbb{R}^n$ as $v = v^+ - v^-$, where $v^+, v^- \in \mathbb{R}^n_+$. Then there is a correspondence between $\mathbb{R}^n$ and $\mathbb{R}^{2n}_+$ where $v \to (v^+, v^-)$. Similarly, a gets sent to $a \to (a^+, a^-)$, and $C \to [C : -C]$.

The transformation from type (ii) to type (iii) is similar. We introduce slack variables $\tilde{v} \in \mathbb{R}^m_+$, and take $v \in \mathbb{R}^n_+$ which is subject to $Cv \leq d$ to $(v, \tilde{v}) \in \mathbb{R}^{n+m}_+$

subject to $C \to [C : I_m]$. To make this all work, we also take $a \to (a, 0)$, where the $0 \in \mathbb{R}^m$.

Lastly, suppose we have a problem of type (iii) and want to convert it to type (i). We take $C \to [C : -C : -I]^T$, $d \to [d : -d : 0]^T$, and $a \to a$.

Lemma. *No interior point of U can be a solution to the linear programming problem unless $a = 0$.*

Theorem. *Consider the linear programming problem of type (iii). If $U \neq \varnothing$, then either $\inf_{v \in U} f(v) = -\infty$ or else the problem has at least one solution.*

Consider a minimizing sequence $u_k \in U$, so that $\lim_{k\to\infty} f(u_k) = \inf_{v\in U} f(v)$. We want to show that $u_k \to u \in U$. Consider the $1 + m \times n$ matrix $B = [a^T, C]^T = [b_1 : b_2 : \ldots : b_n]$. Note that the sequence Bu_k belongs to $\tilde{C} = \{\sum_{i=1}^n v_i b_i : v_i \geq 0\}$, which is convex and closed. Then $Bu_k = [a_0 u_k : Cu_k]^T = [f(u_k) : d]^T \to [\inf_{v\in U} f(v) : d]^T$, which we define as $\beta \in \mathbb{R}^{m+1}$ such that $Bu_k \to \beta$. Since $\tilde{C}$ is closed, $\beta \in \tilde{C}$. Let $u \in \mathbb{R}^n_+$ be such that $\beta = \sum_{i=1}^n u_i b_i$, which is precisely such that $Bu = [a^T, C]^T u = \beta$.

Definition. If U is a convex set of a vector space V, a point $u \in U$ is called an extreme point of U if $u = \lambda v + (1 - \lambda)w$ for $v, w \in U$ implies that $u = v = w$. An extreme point of a polyhedron is called a vertex.

Theorem. *A point $u \in U$ where $u \neq 0$ is a vertex of the polyhedron $U = \{v \in \mathbb{R}^n_+ : \sum_{j=1}^n v_j C^i = d\}$ if and only if the vectors C^j where $u^j \neq 0$ are linearly independent.*

Simplex Method

Suppose we have the problem of finding $u \in U = \{v \in \mathbb{R}^n_+, Cv = d\}$ such that $f(u) = \inf_{v\in U} f(v)$, where $f(v) = (a, v)$. We have previously seen that either the infimum is $-\infty$, or else there is an attainable solution.

For any $v \in U$, we define $I^*(v) = \{j : 1 \leq j \leq n, v_j > 0\}$. If zero is in U, then $I^*(0) = \varnothing$ and it must be a vertex of U.

Theorem. *A nonzero point $u \in U$ is a vertex of $U = \{v \in \mathbb{R}^n_+ : Cv = d\}$ if and only if the vectors C^j for $j \in I^*(u)$ are linearly independent.*

Theorem. *If the problem described has a solution, then at least one vertex is a solution.*

Proof. Let u be a solution. If $u = 0$ is a solution, then it is a vertex. Otherwise, if the C^j columns where $j \in I^*(u)$ are all linearly independent, then u is again a vertex. If the columns are linearly dependent, then there exist w_j (not all zero) such that $w_j = 0$ for $j \neq I^*(u)$ and $\sum w_j C^j = 0$.

Now consider vectors $u + \theta w$, with $\theta \in \mathbb{R}$. Note that $C(u + \theta w) = Cu = d$, so that $u + \theta w \in U$. Furthermore, $I^*(u + \theta w) \subseteq I^*(u)$. Then let $-\infty < \theta_0 = \max\left\{-\frac{u_j}{w_j} : j \in I^*(u),\ w_j > 0\right\} < 0$ and $0 < \theta_1 = \min\left\{-\frac{u_j}{w_j} : j \in I^*(u),\ w_j < 0\right\}$. If we let $\theta_0 < \theta < \theta_1$, then $f(u + \theta w) = (a, u + \theta w) = f(u) + \theta(a, w)$. Since θ can take on negative and positive values, we must have that $(a, w) = 0$, meaning that all vectors of the form $u + \theta w$ are solutions to the original problem.

Now by our definitions, we can consider $u' = u + \theta_0 w$, so that $I^*(u') \subset I^*(u)$. If the C^j for $j \in I^*(u')$ are linearly independent, then we are done since u' is a vertex. If not, we can repeat the process. However, each time we do we get another u'' with the set $I^*(u'') \subset I^*(u')$. Since the cardinality of these sets are decreasing, we must eventually have a set that reduces to nothing. ■

Theorem. *If a polyhedron is nonempty, then it has at least one vertex and the number of vertices is finite.*

These theorems tells us that if there is a solution, then it suffices to check all of the vertices.

Simplex Algorithm

Recall that we were trying to solve the problem of finding $u \in U = \{v \in \mathbb{R}^n_+ : Cv = d\}$ such that $f(u) = \inf_{v \in U} f(v)$, where $f(v) = (a, v)$. We know that if there exists a solution to this problem, then at least one of the vertices of U is a solution.

We now consider the alternative problem of finding $(u, \tilde{u})$ in $\tilde{U} = \{(v, \tilde{v}) \in \mathbb{R}^n_+ \times \mathbb{R}^m_+ : Cv + \tilde{v} = d\}$ such that $\tilde{f}(u, \tilde{u}) = \inf_{(v,\tilde{v}) \in \tilde{U}} \tilde{f}(v, \tilde{v})$, where we define

$\tilde{f}(v,\tilde{v}) = \sum_{i=1}^{m} \tilde{v}_i = (0,\ldots,0,1,\ldots,1)\cdot(v,\tilde{v}) \geq 0$. If v is a solution of the original problem, then $(v,0)$ is a solution to the above problem.

We will now look at the simplex algorithm itself. The idea behind the method is to move from one vertex u_k to another u_{k+1}, with the intent of decreasing the primary function $f(u_{k+1}) < f(u_k)$ which you want to minimize. Recall that a point u_k is a vertex if the columns $\{C^j : j \in I_k\}$ corresponding to the indices of u_k that are nonzero are linearly independent, forming a basis for $\mathbb{R}^m$. The idea is to choose one of these columns C^j to be ejected and introduce another one, so that u_{k+1} is another vertex with $\{C^j : j \in I_{k+1} = (I_k - \{j^-\}) \cup \{j^+\} \notin I_k\}$.

The simplex algorithm involves using elementary row operators to change $Cv = d$ into some $C'v = d'$, where the original set U is preserved. Note that if our minimization problem is $f(v) = a \cdot v$ over $U = \{Cv = d\}$, then it is equivalent to minimize over $f'(v) = (a \pm \theta r)\cdot v$ if r is any row of C.

As an example, suppose we have $U = \{v \in \mathbb{R}^3_+ : 3v_1 - v_2 + 2v_3 \leq 7, -2v_1 + 4v_2 \leq 12, -4v_1 + 3v_2 + 8v_3 \leq 10\}$, and want to minimize $f(v) = v_1 - 3v_2 + 2v_3$. We introduce slack variables $v_4, v_5, v_6 \geq 0$ and begin to build our tableau

$$\begin{array}{cccccc|c} 3 & -1 & 2 & 1 & 0 & 0 & 7 \\ -2 & 4 & 0 & 0 & 1 & 0 & 12 \\ -4 & 3 & 8 & 0 & 0 & 1 & 10 \\ \hline 1 & -3 & 2 & 0 & 0 & 0 & 0 \end{array}$$

Our basic feasible solution is $u_0 = (0,0,0,7,12,10)$. Note that the last entry in the bottom right corner is the value of $f(u_0)$. If all the a_i's are nonnegative, then we stop and take our basic feasible solution, but in this case the second column has a negative a_i. In general, we choose the most negative a_i and denote its column as j^+ (in our case $j^+ = 2$). Suppose that C^{j^+} has at least one positive element, we will choose one of these elements to become a pivot point. The choice relies on comparing each positive element with its d_i value, so we look at $\frac{12}{4}$ and $\frac{10}{3}$, choosing the one that gives us the smallest value.

$$\begin{array}{cccccc|c} \frac{5}{2} & 0 & 2 & 1 & \frac{1}{4} & 0 & 10 \\ -\frac{1}{2} & 1 & 0 & 0 & \frac{1}{4} & 0 & 3 \\ -\frac{5}{2} & 0 & 8 & 0 & -\frac{3}{4} & 1 & 1 \\ \hline -\frac{1}{2} & 0 & 2 & 0 & \frac{3}{4} & 0 & 9 \end{array}$$

Next we choose the top left element to act as a pivot.

$$\begin{array}{cccccc|c} 1 & 0 & \frac{4}{5} & \frac{2}{5} & \frac{1}{10} & 0 & 4 \\ 0 & 1 & \frac{2}{5} & \frac{1}{5} & \frac{3}{10} & 0 & 5 \\ 0 & 0 & 10 & 1 & -\frac{1}{2} & 1 & 11 \\ \hline 0 & 0 & \frac{12}{5} & \frac{1}{5} & \frac{4}{5} & 0 & 11 \end{array}$$

Now we are done, with minimal value -11. Furthermore, $(4, 5, 0, 0, 0, 11)$ is the solution to the larger problem, meaning that $(4, 5, 0)$ is a solution to the original problem.