



Universidade Federal da Paraíba
Centro de Informática
Programa de Pós-Graduação em Modelagem Matemática e Computacional

MODELO DE PLANEJAMENTO DE USINA HIDROTÉRMICA BASEADO EM PROGRAMAÇÃO DINÂMICA DUAL ESTOCÁSTICA

Jefferson Bezerra dos Santos

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional, UFPB, da Universidade Federal da Paraíba, como parte dos requisitos necessários à obtenção do título de Mestre em Modelagem Matemática e Computacional.

Orientadores: Camila Mara Vital Barros
Sérgio de Carvalho Bezerra

João Pessoa
Abril de 2019

Capítulo 1

Despacho de energia

O conjunto de variáveis que identificam o desenvolvido de um país é extremamente complexo. Uma das variáveis utilizada nesse análise é a facilidade que a população tem acesso a estrutura básica de serviços como: saneamento, básico, transportes, as telecomunicações e a energia [1]. Essa última recebe uma atenção especial por ser indispensável para as atividades humanas que necessitam de um algum tipo de maquinário. Por apresentar tal característica o setor energético historicamente é constituído por dois extremos.

O primeiro extremo tem relação com o desenvolvimento tecnológico particularmente a eficiência energética. O intuito da eficiência energética é o desenvolvimento de novas tecnologias que utilizem de forma eficiente a energia disponível, além de possibilitar uma eficiência na forma de produção energética. Para exemplificar o primeiro ponto pode-se considerar o caso do automóvel que por muitas décadas empregava como a sua principal fonte de energia a gasolina, porém gradualmente sua fonte de energia está sendo modificada para atender outras formas de combustível como o etanol [1]. Com a relação ao segundo ponto, na atualidade o principal destaque deve-se as pesquisas por novas fontes de energia como: geotérmica, células de hidrogênio, eólica e solar.

O segundo extremo tem como objetivo garantir que a população tenha acesso a essas novas tecnologias de maneira simples, eficaz e de baixo custo. O principal reflexo dessa iniciativa é observado no fornecimento de energia elétrica [1]. Contudo, sua influência também é percebida em outros setores em uma menor escala. Por exemplo, na década de 70 a substituição da lenha por derivado do petróleo (GLP, gás liquefeito do petróleo)[1]. Com essa mudança um maior número da população teve acesso a uma forma eficiente de energia no tratamento de alimentos. Essa mudança ocorrida foi utilizada como um sinal de modernização do Brasil[1]. A Figura(1.1) a seguir traz um comparativo com as principais diferenças entre a matriz energética brasileira e a matriz mundial.

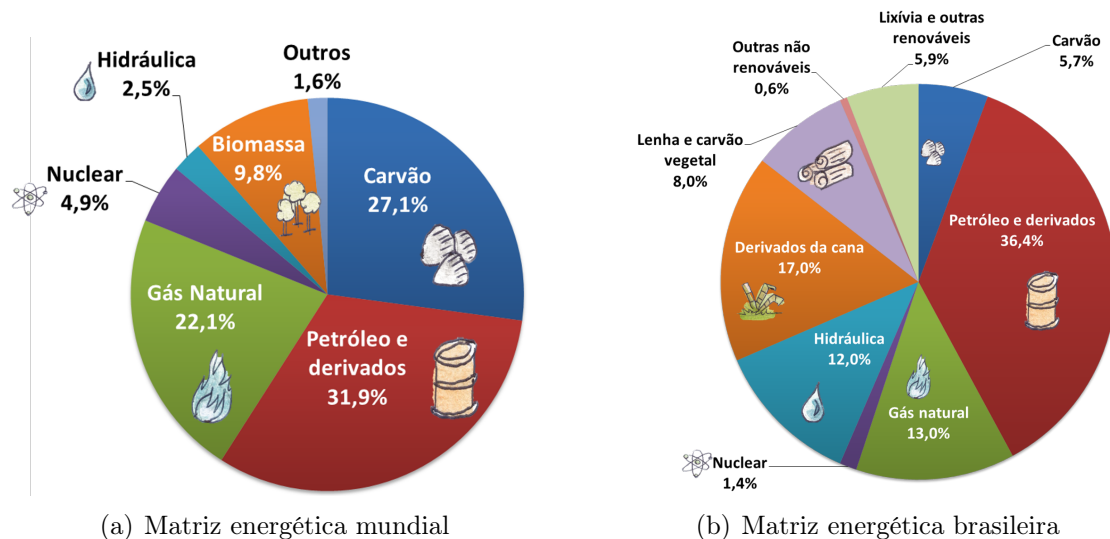


Figura 1.1: Comparação entre a matriz energética mundial e a brasileira.
Fonte: Empresa de Pesquisa Energética.

Nota-se que pela Figura(1.1)(a) que matriz energética mundial é constituída principalmente por fontes não renováveis. Contudo, ao contrário do cenário mundial no Brasil há uma grande utilização de energias renováveis como indica a Figura(1.1)(b) com destaque para a geração elétrica por meio da energia hidráulica.

Dentre todos os segmentos de infra-estrutura o de energia é o mais universalizado [1]. O sistema energético brasileiro abrange uma considerável conjunto de nichos. Contudo, há um quantidade de nichos que constituem uma dificuldade para o sistema. A incidência e as dimensões dos nichos não atendidos estão diretamente relacionada a fatores como a localicazação e as dificuldades físicas ou econômicas para a expansão da rede elétrica [1]. Essas dificuldades devem-se pelo Brasil ser constituído por cinco regiões geográficas: Sul, Sudeste, Norte, Nordeste, Centro-Oeste. Cada uma das regiões mencionadas possui suas características particulares como: relevo, hidrografia e economia. Esses fatores interferem de forma substancial no abastecimento. Uma vez que tais características determinam os contornos que a configuração do sistema de geração e transmissão devem obedecer[1]. Portanto, as características regionais indicam como a população tem acesso a energia elétrica. No caso do Brasil o sistema elétrico é constituído por três grande blocos que constituem o sistema elétrico nacional: Geração, Transmissão e Distribuição. Vale ressaltar que há percas de energia no sistema, isto é, toda energia é utilizada pelo sistema, mas não é possível a utilização de toda a energia disponível. A Figura(1.2) a seguir exemplificar os processos mencionados e as percas de energia que ocorrem no sistema elétrico nacional.

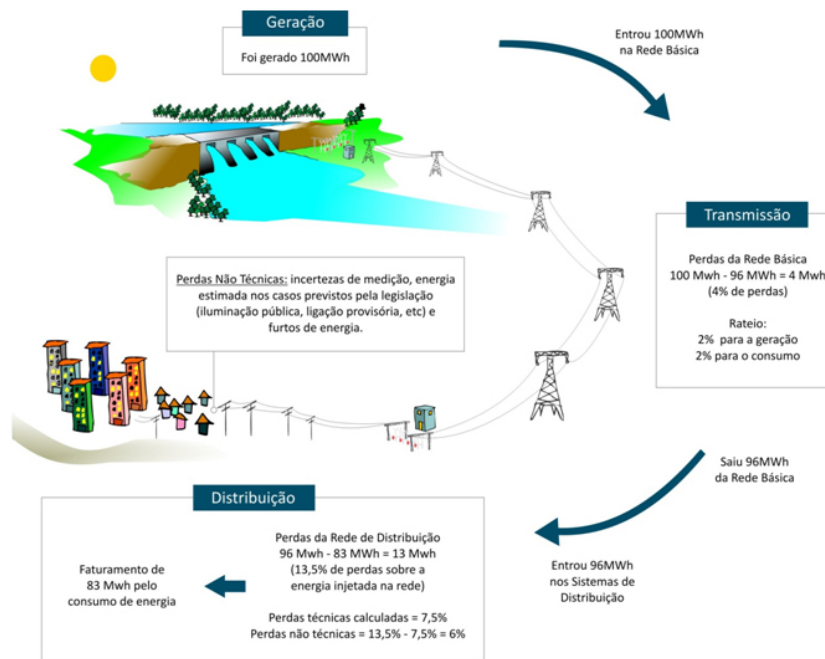


Figura 1.2: Geração, Transmissão e Distribuição.
Fonte: Agência Nacional de Energia Elétrica.

Para o gerenciamento e transmissão da energia elétrica o Brasil possui o Sistema Interligado Nacional (SIN) gerenciado pelo Operador Nacional de Energia (ONS) correspondendo as regiões Sul, Sudeste, Centro-Oeste, Nordeste e parte do Norte. O SIN é responsável por abrigar cerca de 96,6% de toda a capacidade de produção de energia do Brasil, seja por meio de fontes internas de energia ou pela importação de energia como ocorre na usina de Itaipu mediante o controle compartilhado com o Paraguai[1]. A adoção do SIN é justificada tendo por base: o intercâmbio energético, a complementaridade entre fontes de geração de energia e pela sua capacidade de expansão.

O intercâmbio energético permite que regiões que estejam vinculada ao SIN possam auxiliar no suprimento da demanda outras de regiões que por algum fator interno ou externo não conseguem manter a demanda na sua localidade[1]. Por exemplo, considerando-se que determinadas regiões brasileiras podem sofrer com a escassez de chuva o que implica na baixa pluviosidade. Essa região corre o risco de enfrentar problemas de abastecimento caso sua fonte de geração de energia elétrica seja por meio de hidrelétricas. Nesse tipo de situação é totalmente possível que outra região que não esteja enfrentando problemas de escassez auxiliar enviando energia elétrica para atender a localidade que esteja enfretando problemas de abastecimento.

A complementaridade entre fontes de energia tem o mesmo princípio do intercâmbio, contudo seu principal objetivo é permitir que uma ou mais regiões sejam abastecidas por diferente tipos de fonte de energia no intuito do sistema funcionar da melhor forma possível [1]. Isto é, localidades que possuem como sua principal

fonte de energia a hidrelétrica podem sofrer com o baixos índices de seus reservatórios. Nesse tipo de situação é comum a ativação de termelétricas para auxiliar o abastecimento. Esse é um exemplo típico de complementaridade oferecido pelo SIN. Uma vez que a energia elétrica gerada pela hidrelétrica possui um custo menor que a mesma quantidade de energia elétrica produzida por uma termelétrica, portanto, deve-se manter como meta a utilização da hidrelétrica para diminuir possíveis custos ao consumidor[1]. Contudo, apesar da termelétrica possui um custo maior para a geração de energia essa não possui o problema de abastecimento da hidrelétrica. Portanto, a complementaridade para este caso seria configurar a utilização da hidrelétrica para a maioria dos casos e a ativação da termelétrica para auxiliar caso houvesse um pico de demanda ou problemas de abastecimento por fatores como a pluviosidade.

A expansão é caracterizada por permitir que o SIN ao longo do tempo tenha condições de assimilar outras hidrelétricas ou regiões permitindo que o sistema possua condições de garantir a demanda mesmo com o aumento do consumo ou possíveis problemas de abastecimento por outros fatores internos ou externos. Por exemplo, em 2003 o SIN possuía 77,6 mil quilômetros de rede no período de 2008 sua extensão era 89,2 mil quilômetros de rede [1].

O custo associado ao produção de energia é uma das variáveis que influencia o preço. Na análise do sistema o preço varia conforme o tipo de energia utilizada [1]. Desta forma, o planejamento energético depende do custo de produção para determinar o despacho de energia. O despacho de energia é definido como quais usinas devem ser mantidas ativas e quais precisam ser desativadas tomando-se em consideração a demanda, a oferta e o custo de produção do sistema.

Na análise do custo de produção para o despacho de energia outro fator a ser considerado são os sistemas isolados. Os sistemas isolados estão localizados principalmente na região Norte, estados como Amazonas, Roraima, Acre, Amapá e Rondônia. Esta denominação deve-se por não estarem interligados ao SIN e por não permitirem um intercâmbio com outras regiões devido as características geográficas. O funcionamento dos sistemas isolados é predominantemente térmico. Os custos para a geração de energia nesses sistemas são superiores ao SIN por serem predominantemente térmicos e pela sua localização requerer alto custo no transporte de combustíveis [1]. Como alternativa para o barateamento da energia gerada pelos sistemas isolados foi constituído imposto denominado Conta de Consumo de Combustíveis (CCC) que permite subsidiar a compra de combustíveis garantindo que população dessas localidades tenha alguns dos benefícios do SIN.

A produção de eletricidade no sistema brasileiro tem como objetivo principal minimizar os custos de operação e garantir o suprimento de energia em todo o país [7]. Devido ao SIN ser constituído predominantemente por um sistema hidrotér-

mico(hidrelétricas e termelétricas em regime de complementaridade) este é afetado pela incerteza associada a pluviosidade das regiões que o constituem[1]. Além da pluviosidade o sistema hidrotérmico brasileiro é constituído pelas seguinte características:

- *Sazonalidade intra natural.* Além da variabilidade natural ocorre uma variação entre as estações do ano.
- *A complementariedade e diversidade regional.* As bacias brasileiras possuem características físicas e climáticas distintas. Outro ponto a ser observado é que no momento que ocorre uma estiagem no Nordeste as bacias do Sul podem estar com um alto nível dos reservatórios dada a pluviosidade da região, ou seja, há uma complementariedade entre as regiões do Brasil.
- *O acoplamento espacial.* Na forma de cascata as usinas que estão mais perto da jusante possuem dependência de usinas mais perto da montante.
- *O acoplamento temporal.* Na forma de cascata decisões sobre a utilização possuem consequências no futuro.
- *Custo termelétrico.* Usinas termelétricas possuem um custo alto de produção elétrica em relação às hidrelétricas.
- *Aspecto ambiental.* Usinas termelétricas possuem um alto impacto ambiental.

As características mencionadas caracterizam o problema de despacho de energia hidrotérmico. A Figura(1.3) exemplifica o acoplamento temporal e espacial entre as usinas em cascata.

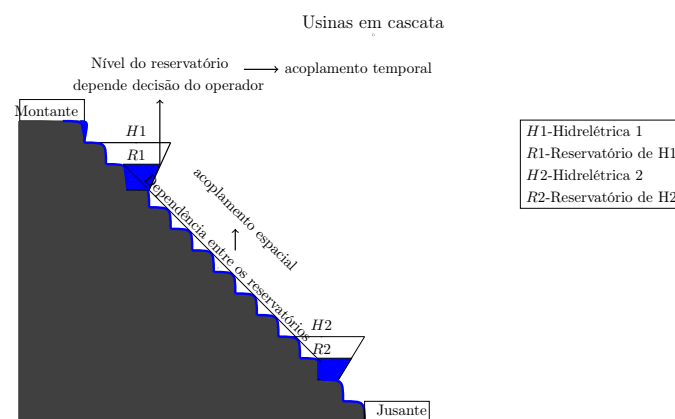


Figura 1.3: Representação do acoplamento espacial e temporal.

Portanto, no planejamento de despacho hidrotérmico a demanda do sistema deve ser garantida de forma a não prejudicar o abastecimento, ao mesmo tempo a geração termelétrica associada aos sistemas isolados possui um custo elevado. Esse

custo deve ser considerado para não ocasionar um aumento desagradável no preço associado ao sistema de energia brasileiro. Nesse contexto, uma definição mais ampla de despacho de energia seria o planejamento eficiente do sistema energético observando características como, demanda, oferta, custo e as configurações do sistema. Para sistemas hidrotérmicos as características do despacho podem ser resumidas no dilema do “operador” dado pelo diagrama a seguir.

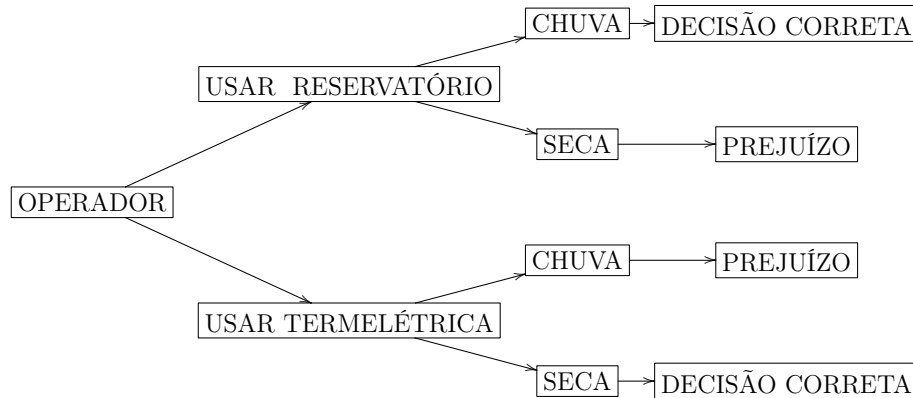


Figura 1.4: Dilema do operador.

Conforme o diagrama na Figura(1.4) o operador do sistema pode ter prejuízo associado a sua escolha dado o componente estocástico relacionado aos reservatórios das hidrelétricas. Pela complexidade do sistema hidrotérmico brasileiro este tipo de decisão possui um grau de dificuldade que transcende a simplicidade. Finalmente, nesse capítulo o problema de despacho de energia foi definido com as suas particularidades. Como as características do sistema energético brasileiro e os principais problemas relacionados as hidrelétricas e termelétricas.

Capítulo 2

Otimização

No planejamento de um sistema energético hidrotérmico o sistema deve preservar as metas de geração para suprir a demanda e minimizar o valor esperado do custo de operação ao longo do período de estudo. As características mencionadas configuram o despacho de energia. Diante das possibilidades de configurações possíveis para o sistema (sequência de acionamento das hidrelétricas e das termoeletricas com o aumento da demanda para uma produtividade fixa) deseja-se obter a combinação na qual o valor do custo associado seja mínimo. Portanto, o despacho de energia trata-se de um problema de otimização. Sejam os conjuntos $D \subset \mathbb{R}^n$ e $\Psi \subset \mathbb{R}^n$ onde $D \subset \mathbb{R}^n$ e uma função $f : \Psi \rightarrow R$ o objetivo é encontrar um minimizador de f no conjunto D [4]. Isto é,

$$\min_{x \in D} f(x), \quad (2.1)$$

define-se o conjunto D como o conjunto viável do problema, os pontos de D serão chamados pontos viáveis e f será chamada de função objetivo. Uma vez formulado o problema em Eq.(2.1) para a sua resolução é necessário uma definição dos elementos que constituem um problema de otimização. Nesta perspectiva são abordadas as definições consideradas de relevância para o presente estudo dadas a seguir.

Definição 2.0.1. Um ponto $\bar{x} \in D$ é,

1. minimizador global para Eq.(2.1) se,

$$f(\bar{x}) \leq f(x), \forall x \in D;$$

2. minimizador local da Eq.(2.1), se existe uma vizinhança U de \bar{x} tal que

$$f(\bar{x}) \leq f(x), \forall x \in D \cap U.$$

Pode-se definir de maneira equivalente o minimizador local como,

$$\begin{aligned} \bar{x} \in D \text{ é minimizador local} \\ \Updownarrow \\ \exists \epsilon > 0 \text{ t.q. } f(\bar{x}) \leq f(x), \forall x \in \{x \in D; \|x - \bar{x}\| \leq \epsilon\} \end{aligned} \quad (2.2)$$

Adotando-se x_g e x_l minimizador global e local respectivamente o entendimento intuitivo das ideias apresentadas é dado pela Figura(2) a seguir:

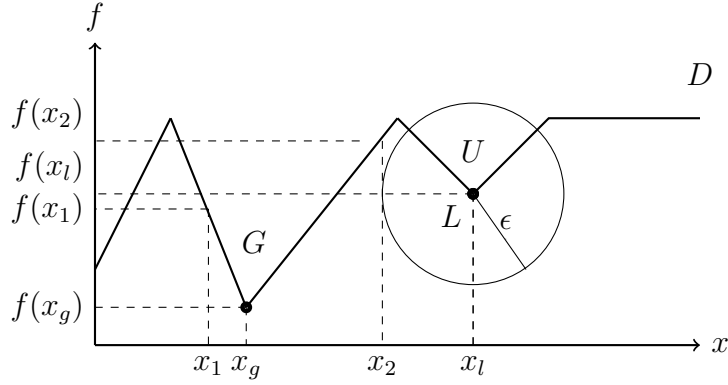


Figura 2.1: Minimizador global (x_g), local (x_l) e vizinhança (U).

O significado da formulação apresentada na Figura(2) é que um minimizador local existe quando é possível encontrar uma região U em D na qual seja satisfeita a relação $f(\bar{x}) \leq f(x)$. Nota-se pela definição 2.0.1 que um minimizador global também é um minimizador local, contudo a recíproca não é verdadeira [4]. O fato do minimizador local necessitar de apenas uma vizinhança U em D indica a possibilidade da função f ter vários minimizadores locais no conjunto viável D . É importante salientar que dependendo do problema de otimização pode existir restrições no conjunto viável D .

Considerando-se o problema,

$$\min_{x, x_d \in D; x \geq x_d} f(x) \quad (2.3)$$

a situação é exemplificada em Figura(2.2).

De acordo com Figura(2.2) apesar de x_g ser o mínimo de f a restrição $x \geq x_d$ torna inviável x_g ser o minimizador para a função f considerando-se o conjunto viável D , portanto, neste caso o minimizador corresponde ao elemento x_l .

Considerando-se o problema de minimização na Eq.(2.1) o valor ótimo é o menor valor que $f(x)$ assumir considerando-se $x \in D$ essa ideia embasar a definição 2.0.2.

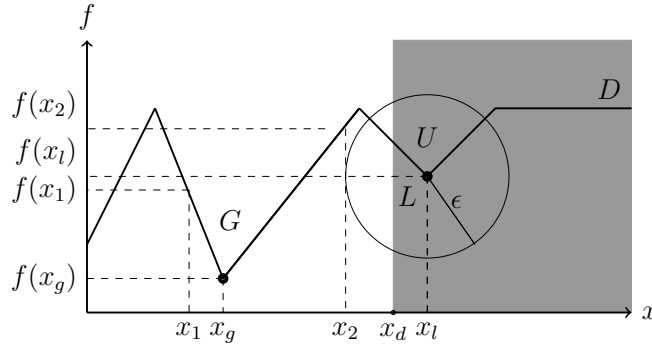


Figura 2.2: Mínimo global (x_g), local (x_l) e conjunto com restrições D .

Definição 2.0.2. Seja $\bar{v} \in [-\infty, +\infty)$ definido por,

$$\bar{v} = \inf_{x \in D} f(x)$$

\bar{v} é o valor ótimo da Eq.(2.1).

Nota-se pela definição 2.0.2 que o objeto a ser encontrado (valor ótimo) está bem definido. O fato de tomar o \inf indica que o valor ótimo pode não pertence ao conjunto de restrições D , contudo deve estar suficiente próximo para ser considerado de relevância para não comprometer a solução do problema dado pela 2.1. De forma geral problemas de minimização estão relacionados com problemas de maximização. De fato, uma relação bastante conhecida entre problemas de maximização e minimização é encontrada em [4] dada por,

$$\max_{x \in D} f(x) \Leftrightarrow \min_{x \in D} -f(x). \quad (2.4)$$

Desta forma, problemas de minimização e maximização podem ser interpretados como o mesmo tipo de problema, entretanto as soluções locais e globais do problema 2.4 evidentemente possuem sinais opostos. A exemplificação deste fato é dada pela Figura(2.3).

O conjunto viável de um problema de otimização é definido por um conjunto de igualdades e/ou desigualdades e/ou inclusão da seguinte forma,

$$D = \left\{ x \in \Psi; \begin{array}{l} h_i(x) = 0, i = 1, \dots, l \\ g_j(x) \leq 0, j = 1, \dots, m \end{array} \right\}$$

ou de forma equivalente,

$$D = \{x \in \Psi; h(x) = 0, g(x) \leq 0\}.$$

dado que $\Psi \subset \mathbb{R}^n$, $h : \Psi \rightarrow \mathbb{R}^l$ e $g : \Psi \rightarrow \mathbb{R}^m$ admitindo-se que f está definida

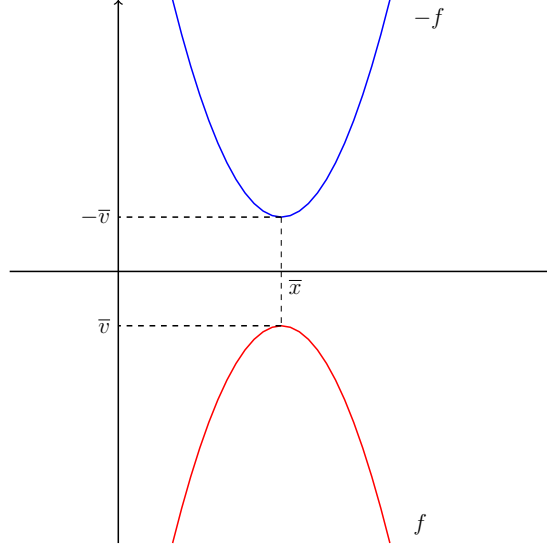


Figura 2.3: Minimização ($-f$) e Maximização (f).

no conjunto Ψ . O conjunto Ψ representa as condições diretas e as desigualdades são conhecidas como restrições funcionais. Na prática a separação entre restrições diretas e funcionais é apenas uma conveniência. De fato, seja a condição direta $x \in \Psi = \mathbb{R}_+^n$ é equivalente a representação funcional dada por $x \in \{x \in \mathbb{R}^n; x_i \geq 0, i = 1, \dots, n\}$ [4]. No conjunto de restrições D em alguns casos existe a necessidade de manipulações e/ou transformações para tornar possível ou facilitar a resolução do problema de otimização. Por exemplo, as restrições de igualdade $h_i = 0$ podem ser reescritas de maneira equivalente como inequações, ou seja,

$$\begin{aligned} h_i(x) &= 0 \\ \Updownarrow & \quad \text{com } i = 0, 1, \dots, l \\ h_i &\leq 0, -h_i(x) \leq 0. \end{aligned}$$

De forma análoga pode-se transformar restrições de desigualdade mediante uma transformação inserindo nas inequações termos artificiais conhecidos como variáveis de folga, isto é,

$$\begin{aligned} g_i(x) &\leq 0 \\ \Updownarrow & \quad \text{com } i = 0, 1, \dots, l \\ g_i + z_i^2 &= 0, z_i \in \mathbb{R}. \end{aligned}$$

Contudo, na prática dependendo do contexto o problema de otimização pode ser tornar mais complexo com tais transformações sua utilização dever ocorrer quando realmente existe alguma indicação do melhoramento do problema transformado em

comparação com o problema original. Em primeiro lugar existe a possibilidade do problema relaxado possui condições de otimalidade mais fracas que o problema com as restrições originais [4]. Outro fator a ser considerado é o aumento no número de restrições. Quanto ao tipo de problema em relação ao conjunto de restrições (D), por enquanto deve-se considerar basicamente dois casos específicos: sem restrições quando $D = R^n$, isto é, sem restrições no domínio no qual f é definida. Uma exemplificação deste caso pode ser observado em Figura(2). Para o caso $D \neq R^n$ tem-se um problema com restrições, ou seja, deve-se considerar um subconjunto do domínio de f um exemplo é dado em Figura(2.2). O conjunto das restrições é de fundamental importância para o estudo da otimização [4]. De fato, uma classe importante de problemas é definida quando o conjunto das restrições forma um poliedro.

Definição 2.0.3. Um conjunto é dito poliedral quando ele pode ser representado como um conjunto de soluções de um sistema finito de equações e inequações lineares, isto é,

$$D = \{x \in R^n; Ax = a, Bx \leq b\},$$

onde:

- $A \in R(l, n)$, $B \in R(m, n)$;
- $a \in R^l$ $b \in R$.

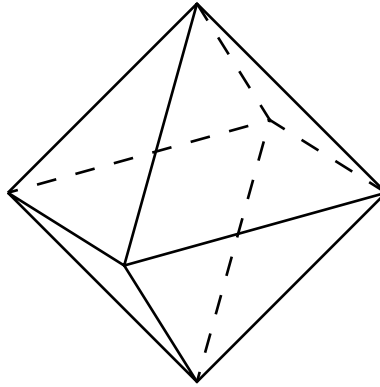


Figura 2.4: Exemplo de um conjunto poliedral \mathbb{R}^3 .

Desta forma tomando-se $h : R^n \rightarrow R^l$ e $g : R^n \rightarrow R^m$ definidas como $h(x) = Ax - a$ e $g(x) = Bx - b$ funções afins nesse contexto [4]. Nota-se que para o caso em \mathbb{R}^2 teríamos um polígono no plano. Uma vez que o conjunto poliedral está bem definido pode-se estabelecer informações úteis sobre a função objetivo f .

Definição 2.0.4. Uma função $f : R^n \rightarrow R$, definida por,

$$f(x) = \langle Qx, x \rangle + \langle q, x \rangle$$

onde $Q \in R(n, n)$, $q \in R^n$, é chamada função quadrática.

Dada Q em 2.0.4 pode-se admitir sem perda de generalidade que está é simétrica, portanto,

$$\langle Qx, x \rangle = \frac{1}{2} \langle (Q + Q^T)x, x \rangle.$$

Uma vez que é possível a substituição da matriz Q pela simétrica $\frac{(Q+Q^t)}{2}$ [4]. Desta forma considera-se dois casos de grande importância: o conjunto (D) é poliedral e a função (f) é quadrática trata-se de um problema de programação quadrático para o caso que o conjunto (D) é poliedral e a função objetivo (f) seja linear, ou seja, quando $Q = 0$ em 2.0.4 trata-se de um problema de programação linear. Para o presente trabalho será considerado somente problemas de programação linear. Contudo, somente a informação sobre o conjunto D ser um conjunto poliedral não permite considerações que facilitem a busca pela a solução do problema liner, além de não permite verificar se a solução encontrada é realmente adequada para o problema proposto na Eq.(2.1). Com esse intuito será tratado somente os aspectos de relevância da análise convexa para o tratamento adequado do conjunto poliedral (D) e para o desenvolvimento de outros resultados importantes para o presente estudo.

Definição 2.0.5. Um conjunto $Q \subset R^n$ é chamado conjunto convexo se para quaisquer $x \in Q$, $y \in Q$ e $\forall \lambda \in [0, 1]$ tem-se que $\lambda x + (1 - \lambda)y \in Q$. Onde tem-se que o ponto $\lambda x + (1 - \lambda)y$ é uma combinação convexa de x e y com o parâmetro λ .

Dada a definição 2.0.5 para que um conjunto Q qualquer seja convexo é necessário que dados dois pontos de Q o segmento de reta cujos os extremos sejam esses pontos deve está no conjunto Q .

Definição 2.0.6. Se $Q \subset R^n$ é um conjunto convexo, diz-se que a função $f : Q \rightarrow R$ é convexa em Q quando para quaisquer $x \in Q$, $y \in Q$ e $\lambda \in [0, 1]$ tem-se que,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

As definições 2.0.5 e 2.0.6 caracterizam os elementos principais de convexidade por meio da convexidade é possível determinar características favoráveis na resolução da Eq.(2.1). Por exemplo, os resultados a seguir permitem descrever os objetos que formam o conjunto viável (D) como conjuntos convexas.

Proposição 2.0.1. Qualquer semi-espço em \mathbb{R}^n , isto é, o conjunto da forma $D = \{x \in R^n; \langle x, a \rangle \leq c\}$ onde $a \in R^n$ $c \in R$ é convexo.

Prova: De fato, tomando-se $x, y \in D$ e definindo uma $f : D \rightarrow R$ tal que,

$$f(x) = \langle x, a \rangle = \sum_{i=1}^n x_i a_i \text{ com } x, a \in D, \text{ onde } a \text{ é fixo e } i = 1, 2 \dots n.$$

Portanto, da hipótese tem-se que

$$\langle x, a \rangle = \sum_{i=1}^n x_i a_i \leq c \quad (2.5)$$

$$\langle y, a \rangle = \sum_{i=1}^n y_i a_i \leq c. \quad (2.6)$$

Tomando $\alpha \in [0, 1]$ e multiplicando 2.5 por (α) e 2.6 por $(1 - \alpha)$ tem-se,

$$\langle \alpha x, a \rangle = \alpha \sum_{i=1}^n x_i a_i \leq \alpha c \quad (2.7)$$

$$\langle \alpha y, a \rangle = (1 - \alpha) \sum_{i=1}^n y_i a_i \leq (1 - \alpha)c. \quad (2.8)$$

Somando-se 2.7 com 2.8 segue que

$$\begin{aligned} \langle \alpha x, a \rangle + \langle (1 - \alpha)y, a \rangle &= \alpha \sum_{i=1}^n x_i a_i + (1 - \alpha) \sum_{i=1}^n y_i a_i \leq \alpha c + (1 - \alpha)c \Rightarrow \\ \langle \alpha x, a \rangle + \langle (1 - \alpha)y, a \rangle &= \alpha f(x) + (1 - \alpha)f(y) \leq c \end{aligned}$$

pela definição 2.0.5 a prova está terminada.

Proposição 2.0.2. Qualquer hiperplano em \mathbb{R}^n , isto é, o conjunto da forma $D = \{x \in \mathbb{R}^n; \langle x, a \rangle \leq c\}$ onde $a \in \mathbb{R}^n$ $c \in \mathbb{R}$ é convexo.

Prova:

De maneira análoga a proposição 2.0.1 deve-se utilizar "=" ao invés de " \leq ".

Proposição 2.0.3. Sejam $D_i \subset \mathbb{R}^n$ tal que $i \in I$ conjuntos convexos onde I é um conjunto qualquer (possivelmente infinito).

Então a intersecção $D \cap_{i \in I} D_i$ também é um conjunto convexo [4].

Corolário 2.0.1. Um conjunto poliedral em \mathbb{R}^n é convexo [4].

Prova:

De fato, o conjunto viável de Eq.(2.1) é um poliedro formado pela intersecção de semi-espacos e hiperplanos nota-se pelas preposições 2.0.1 e 2.0.2 que tais conjuntos são convexos, logo, segue pela proposição 2.0.3 que conjunto (D) é convexo.

Desta forma, o conjunto problemas de otimização quadráticos e lineares com o conjunto (D) de restrições podem sem perda de generalidade serem tratados como conjuntos convexos, portanto, quando houve referência ao conjunto viável (D) será considerado tanto como um conjunto poliedral como também um conjunto convexo. A proposição 2.0.3 e o corolário 2.0.1 podem ser encontrados com maiores detalhes

em [4]. O fato de pode-se trabalhar com o conjunto das restrições de um problema de otimização como um conjunto convexo deve-se ao Teorema de Maximização de uma função convexa em um conjunto poliedral diretamente deste teorema obtém-se uma forma de assegurar a existência de uma solução para um problema de programação linear. Além de indicar uma formulação para a busca da solução.

Teorema 2.0.1. Sejam $D \subset \mathbb{R}^n$ um conjunto poliedral que não contém nenhuma reta e $f : D \rightarrow \mathbb{R}$ uma função convexa. Supondo-se que o problema dado por

$$\max_{x \in D} f(x) \quad (2.9)$$

possua uma solução. Então existe uma solução deste problema que é um vértice de D [4].

Conforme o teorema 2.0.1 desde que as suposições sejam respeitadas tem-se uma maneira de encontrar uma solução para o problema 2.9. Nota-se que o problema tem características que se assemelham ao problema original dado por 2.1. Contudo, a forma apresentada no teorema 2.0.1 ainda não é a procurada. O resultado do corolário 2.0.2 estabelece a relação final que o problema esteja bem definido, além de permite um melhor análise do problema proposto por 2.1.

Corolário 2.0.2. Supondo-se que $D \subset \mathbb{R}^n$ seja um conjunto poliedral que não contém nenhuma reta e que o problema de programação linear

$$\min \langle c, x \rangle \text{ sujeito a } x \in D \quad (2.10)$$

onde $c \in \mathbb{R}^n$, tenha uma solução [4].

Demonstração:

De fato, aplicando a relação entre problemas de maximização e minimização abordada em 2.4 o problema 2.10 é reescrito como,

$$- \max \langle -c, x \rangle \text{ sujeito a } x \in D.$$

a função objetivo desse problema é convexa, portanto, o resultado segue da aplicação do teorema 2.0.1. Para que as condições do teorema sejam satisfeitas é necessária uma restrição no conjunto viável D , portanto, tomando $x \geq 0$ garante que o conjunto viável D não possui nenhuma reta o que permite que seja possível utilizar o resultado.

Em problemas de natureza prática é comum que a variável de interesse x assumir valores não negativos, consequentemente a restrição $x \geq 0$ não se torna um empecilho para a maioria das aplicações. Por exemplo, considerando-se que uma determinada empresa entregar três tipos de mercadoria A, B e C. Cada mercadoria

tem um custo c_1, c_2 e c_3 associado que depende do tempo de entrega t_1, t_2 e t_3 respectivamente. Considerando-se que o custo total possui uma relação com o tempo de tal forma que quanto menor o tempo de entregar menor custo associado. Contudo, deve-se garantir a integridade de cada mercadoria com base em características específicas. Portanto, este tipo de problema poderia ser modelado como,

$$\min_{x \in D} \langle c, t \rangle$$

onde:

- c é o vetor custo;
- t é o vetor tempo;
- D conjunto viável com as restrições de qualidade.

Nota-se que nesse tipo de situação não há um sentido para que a variável de interesse t seja considerada negativa. Desta forma, para a utilização do resultado do corolário 2.0.2 tem-se o conjunto viável D dado por,

$$D = \{x \in R_+^n; Ax = a, Bx \leq b\}.$$

Observa-se que o resultado apresentado pelo corolário 2.0.2 não garante que o conjunto D seja limitado, portanto há a possibilidade que o conjunto viável do problema de programação linear não seja limitado nesse caso pode ocorrer problemas quanto a buscar pela solução. A resolução de um problema de programação linear de maneira direta, ou seja, buscando-se o minimizador da função objetivo f no conjunto viável D pode não ser a melhor alternativa a depender do caso. Nesta perspectiva é constituído o conceito de dualidade de um problema de programação linear. O conceito de dualidade a grosso modo é uma forma de transformação de um problema original em um outro problema. Ambos os problemas possuem a mesma solução desde que determinadas características sejam satisfeitas. A principal justificativa para a utilização do problema auxiliar deve-se ao fato que caso o problema original possua um tratamento complexo frequentemente o problema auxiliar possui uma menor complexidade.

Definição 2.0.7. Considerando-se o seguinte problema de programação de linear que será denominado de problema primal,

$$\min \langle c, x \rangle \text{ sujeito a } x \in D = \{x \in R^n; Bx \geq b\}, \quad (2.11)$$

onde $b \in R^m$, $c \in R^n$ e $B \in R(m, n)$.

O problema dual de 2.11 é definido como,

$$\max \langle b, \mu \rangle \text{ sujeito a } \mu \in \Delta = \{\mu \in R_+^m; B^T \mu \leq c\} \quad (2.12)$$

Algumas considerações que devem ser tomadas no contexto da dualidade de um problema primal. Primeiramente, o problema primal é constituído n variáveis com m restrições, por outro lado para o problema dual são m variáveis com n restrições. A natureza dos resultados para a dualidade não ser modificar mesmo que o formato do problema primal seja modificado, porém, esta afirmação é verdadeira deste que o problema dual esteja definido adequadamente. Por fim, o seguinte resultado garante a relação de equivalência entre o problema primal e dual.

Teorema 2.0.2. Dualidade em programação linear.[4]

Para o par de problemas 2.11 e 2.12 a seguinte três possibilidades são possíveis:

- (1) Ambos os problemas possuem o conjunto viável vazio, i.e, $D = \emptyset$ e $\Delta = \emptyset$.
- (2) A função objetivo de um dos problemas é limitada (inferiormente para 2.11 e superiormente para 2.12) no conjunto viável correspondente não-vazio. Neste caso, o outro problema tem o seu conjunto viável vazio.
- (1) A função objetivo de um problemas é limitada (inferiormente para 2.11 e superiormente para 2.12) no conjunto viável correspondente não-vazio. Neste caso, ambos problemas possuem solução e seus valores ótimos coincidem.

A terceira possibilidade indica que caso seja possível a solução do problema primal ou dual o valor ótimo da função objetivo de ambos os problemas coincide, ou seja, pode-se resolver o problema primal de forma indireta baseando-se no dual associado. Nesse capítulo foram abordados os principais aspectos da teoria da otimização relevante para o estudo sobre o problema do despacho de energia. Desde a definição do conjunto viável e as relações que podem ser estabelecidas, além dos principais conceitos para a resolução de problemas de programação linear como a teoria da dualidade.

Capítulo 3

Probabilidade

No capítulo anterior estabeleu-se os conceitos da teoria da otimização relevantes para essa pesquisa em particular problemas de programação linear. Contudo, para a formulação de um modelo de despacho de energia considerando somente o aspecto determinístico, isto é, um problema de programação linear esse não possui as características necessárias para a análise do sistema hidrotérmico, pois, as mudanças que podem ocorrer no ambiente acarretam um sério impacto no sistema. Portanto, para modelos hidrotérmicos é necessário o conhecimento sobre a teoria das probabilidades. Este capítulo aborda os principais conceitos para o entendimento do modelo de despacho de energia considerando a variedade de cenários que podem ocorrer. Para um estudo mais rigoroso sobre a teoria da probabilidade consulte [5] e [6].

Na natureza encontramos uma série de situações que envolvem algum tipo de incertezas denominados fenômenos ou experimentos aleatórios. Por exemplo, o lançamento de um dado, a previsão do clima e o lançamento de um moeda são exemplos típicos. O espaço amostral é o conjunto de todos os resultados possíveis sendo representado por Ω podendo ser enumerável, finito ou infinito [6]. Neste capítulo será adotado as seguintes convenções:

- ω é um elemento do espaço amostral Ω ;
- A utilização de letras maiúsculas para a nomeação de subconjuntos, por exemplo $A \subset \Omega$;
- O conjunto vazio será denotado por \emptyset .

Um exemplo simples da utilização do espaço amostral Ω é o lançamento de um dado. Seja o lançamento de um dado desejando-se uma descrição do espaço amostral desse evento deve-se elencar todos os resultados possíveis, ou seja, 1, 2, 3, 4, 5, 6. Portanto, o espaço amostral será dado por,

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Alguns dos subconjuntos de Ω podem ser elencados da seguinte forma,

$$A = \{1, 2, 3\}, B = \{6\}.$$

A representação pela linguagem dos conjuntos é usualmente empregada em probabilidade com o intuito de uma descrição objetiva dos eventos. Por exemplo, considerando-se a descrição do evento “sair um número menor que 4” na linguagem dos conjuntos a representação desse evento poderia ser,

$$C = \{\omega \in \Omega; \omega < 4\}$$

desta forma, o evento de interesse está bem formulado. Na modelagem de um evento probabilístico por meio da linguagem dos conjuntos deve-se tomar alguns cuidados básicos, por exemplo,

$$6 \in \Omega, 6 \subset \Omega.$$

Nota-se que a segunda afirmação não possui sentido algum. As principais notações para as operações entre conjuntos são:

- (i) A^c complemento, isto é, todos os elementos ω de Ω exceto aqueles que estão em A .
- (ii) $A_1 \cup A_2 \cup A_3, \dots, \cup A_n$ ou $\bigcup_{j=1}^n A_n$, é a união, são todos os pontos $\omega \in \Omega$, que pertencem a pelo menos um A_i com $i = 1, 2, \dots, n$.
- (iii) $A_1 \cap A_2 \cap A_3, \dots, \cap A_n$ ou $\bigcap_{j=1}^n A_n$, é a intersecção, são todos os pontos $\omega \in \Omega$, pertencentes A_i qualquer que seja o $i = 1, \dots, n$.
- (iv) $A - B$ ou $A \cup B^c$ é a diferença entre A e B , ou seja, são todos os elementos $\omega \in \Omega$ pertencentes à A que não pertencem a B .
- (v) $A \triangle B$ é a diferença simétrica entre A e B , todos os $\omega \in \Omega$ pertencentes a $A \cup B$ exceto os $\omega \in \Omega$ que estão na $A \cap B$.

As notações de operações permitem estabelecer uma série de análises dos eventos de interesse do espaço amostral Ω . Considerando-se dois conjuntos A e B são ditos disjuntos ou mutuamente exclusivos se, somente se, a intersecção entre A e B é o conjunto vazio, ou de outra forma,

$$A \text{ e } B \text{ disjuntos} \Leftrightarrow A \cap B = \emptyset. \quad (3.1)$$

Nota-se que a relação estabelecida em 3.1 é uma forma de descrever os eventos do espaço amostral Ω que não possuem nenhum elemento ω em comum. Pode-se

particionar o conjunto Ω por subconjuntos desde que os subconjuntos escolhidos de Ω respeitem os dois critérios a seguir,

$$\bigcup A_i = \Omega \quad e \quad A_i \cap A_j = \emptyset, i \neq j.$$

Uma vez que os principais conceitos da linguagem dos conjuntos foram estabelecidos pode-se considerar a definição clássica de probabilidade [5].

Definição 3.0.1. Define-se a probabilidade clássica de um evento A ocorre pela razão entre o números de casos favoráveis pelo número de casos possíveis, ou seja,

$$P(A) = \frac{n(A)}{n(\Omega)}. \quad (3.2)$$

A relação estabelecida em 3.0.1 é válida para um Ω finito. Além disso, os eventos de Ω devem ser equiprováveis, isto é, cada evento tem a mesma probabilidade de ocorrência. A definição 3.0.1 segue o princípio da indiferência, ou seja, a ocorrência de um evento A não modificar a probabilidade de ocorrência de um evento B [5]. Para o lançamento do dado tem-se,

$$P(i) = \frac{1}{6}, \forall \quad i \in \Omega, \quad i = 1, 2, 3, 4, 5, 6.$$

Portanto, independente da jogada e considerando-se que o dado não é viciado há uma mesma probabilidade associada independente do evento no espaço amostral. Outro exemplo clássico é o lançamento de uma moeda. Descrevendo os eventos possíveis por,

$$\Omega = \{\text{cara}, \text{coroa}\}.$$

Nota-se ambos os eventos são equiprováveis, ou seja,

$$P(\text{cara}) = P(\text{coroa}) = \frac{1}{2}.$$

Uma outra definição usual de probabilidade é a frequentista ou estatística.

Definição 3.0.2. Dado um evento A definindo-se $n(A)$ o número de ocorrências independentes do evento A . Seja $n(\Omega)$ todos os casos possíveis. Define-se a probabilidade frequentista como,

$$P(A) = \lim_{n \rightarrow \infty} \frac{n(A)}{n(\Omega)}.$$

A expressão em 3.0.2 é a frêcia da ocorrência do evento A para o n suicientemente grande. As definições de probabilidade mencionadas no até o presente momento utilizam-se um forte apelo a intuição. Com intuito de uma definição mais rigorosa de probabilidade o matemático Kolmogorov estabeleceu um conjunto de axiomas [6]. A seguir define-se uma probabilidade de forma geral.

Definição 3.0.3 (Probabilidade). Uma função P definida na σ -álgebra F de subconjunto de Ω para valores no intervalo $[0,1]$, é uma probabilidade se satisfaz os Axiomas de Kolmogorov:

1. $P(\Omega) = 1$;
2. Para todo subconjunto $A \in F$, $P(A) \geq 0$;
3. Qualquer que seja a sequência $A_1, A_2, \dots \in F$ mutuamente exclusivos tem-se,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

O axioma (1) indica que uma probabilidade não assume valores maiores que 1. Além disso, pelo axioma(2) qualquer que seja a probabilidade essa é não negativa. Os axiomas (1) e (2) garatem que uma probabilidade qualquer pertence ao intervalo $[0, 1]$. Finalmente, o axioma (3) permite a decomposição da probabilidade da união para eventos mutuamente exclusivos, ou seja, eventos independentes. A trinca (Ω, F, P) é denominado espaço da probabilidade os subconjuntos em F são os eventos, somente a estes tem-se uma probabilidade associada. A definição de probabilidade em 3.0.3 permite a caracterização de propriedades com os eventos de Ω . Primeiramente, considere o espaço de probabilidade (Ω, F, P) onde os conjuntos mencionados estão contidos neste espaço tem-se:

1. $P(A) = 1 - P(A^c)$
2. Sendo A e B dois eventos quaisquer tem-se:

$$P(B) = P(B \cap A) + P(B \cap A^c)$$
3. Se $A \subset B$, então $P(A) \leq P(B)$
4. Regra da adição de probabilidade

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
5. Para eventos quaisquer A_1, A_2, \dots

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i).$$

Para exemplificar as propriedades apresentadas seja um dado equilibrado esse é lançado duas vezes e as faces resultantes observadas [6]. Considerando-se os seguintes eventos de interesse:

A : a soma dos resultados é ímpar.

B : o resultado do primeiro lançamento ímpar.

C : o produto do resultado é ímpar.

Primeiramente deve-se definir um espaço adequado para o experimento, considerando-se $\Omega = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\}$. O conjunto Ω é o produto cartesiano de todos os resultados possíveis, ou seja, todo ponto $\omega \in \Omega$ pode ser descrito como $\omega = (\omega_1, \omega_2)$. Utilizando-se como σ -álgebra o conjunto das partes de Ω e P a probabilidade associada a cada ponto para cada um dos eventos em Ω tem-se uma probabilidade uniforme, ou seja, $P(\{\omega\}) = \frac{1}{36}$. De fato, o resultado é uma implicação do princípio fundamental da contagem. Se uma tarefa tem k etapas e cada etapa i tem n maneiras diferentes de ser realizada, então o número total de alternativas para realizar a tarefa é o produto $n_1 n_2 \dots n_k$ [6]. Para cada jogada de um dos dado existe 6 possibilidades para o resultado. Portanto, para o lançamento de dois dados tem-se 36 possibilidades significando que o espaço amostral possui 36 pontos. A jogada do dado D_1 não interfere na jogada do dado D_2 , ou seja, existe independência de eventos. Logo, o evento A pode ser descrito como,

$$A = \{\omega = (\omega_1, \omega_2) \in \Omega; \omega_1 + \omega_2 \text{ é ímpar}\}$$

ou de maneira equivalente,

$$A = \{\omega = (\omega_1, \omega_2) \in \Omega; \omega_1 + \omega_2 = 2n + 1, n \in \mathbb{N}\}.$$

Para o cálculo da probabilidade do evento poderia-se elencar todos as possíveis combinações para este evento, ou seja, os pontos procurados são $\{1, 2\}, \{1, 4\}, \{1, 6\}, \{2, 1\}, \{2, 3\}, \{2, 5\}, \{3, 1\}, \{3, 2\}, \{3, 4\}, \{3, 5\}, \{3, 6\}, \{4, 1\}, \{4, 3\}, \{4, 5\}, \{5, 1\}, \{5, 2\}, \{5, 4\}, \{5, 6\}, \{6, 1\}, \{6, 3\}, \{6, 5\}$. Verifica-se que o total é de 18 pontos pontos a probabilidade associada é $P(A) = \frac{18}{36} = \frac{1}{2}$. De forma análoga o evento B pode ser descrito da seguinte forma,

$$B = \{\omega = (\omega_1, \omega_2) \in \Omega; \omega_1 \text{ é ímpar}\}.$$

Os pontos são $\{1, 1\}, \{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{1, 6\}, \{3, 1\}, \{3, 2\}, \{3, 3\}, \{3, 4\}, \{3, 5\}, \{3, 6\}, \{5, 1\}, \{5, 2\}, \{5, 3\}, \{5, 4\}, \{5, 5\}, \{5, 6\}, \{6, 1\}, \{6, 2\}, \{6, 3\}, \{6, 4\}, \{6, 5\}, \{6, 6\}$. A probabilidade de ocorrência do evento B é $P(B) = \frac{18}{36} = \frac{1}{2}$. Finalmente, a probabilidade associada a C é $P(C) = \frac{9}{36} = \frac{1}{4}$. Com base nas informações das probabilidade do eventos A, B e C pode-se fazer conclusões interessantes, por exemplo, percebe-se que $P(A \cap B) = \frac{1}{4}$. De fato, $A \cap B = \{1, 1\}, \{1, 3\}, \{1, 5\}, \{3, 1\}, \{3, 3\}, \{3, 5\}, \{5, 1\}, \{5, 3\}, \{5, 5\}$.

segue que,

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \\ P(A \cup B) &= \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4}. \end{aligned}$$

Consequentemente, a probabilidade da soma ser ímpar e o resultado do primeiro lançamento ser ímpar corresponde a $\frac{1}{4}$. E a probabilidade da ocorrência da soma dos resultados ser ímpar ou resultado do primeiro lançamento ser ímpar corresponde a $\frac{3}{4}$. Geralmente em problemas de natureza prática há a necessidade do cálculo da probabilidade um certo A dado que o um evento b ocorreu. Por exemplo, perguntas do tipo "Qual a probabilidade que ocorra uma chuva dado que a temperatura diminuiu". Esse tipo de pergunta onde o interesse de um evento depende da ocorrência de outro evento fundamenta a noção de probabilidade condicional.

Definição 3.0.4 (Probabilidade Condicional). Considerando-se os eventos A e B em (Ω, \mathcal{F}, P) onde $P(B) > 0$. A probabilidade de ocorrência do evento A tal que B ocorreu é dado por,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

para o caso $P(B) = 0$ define-se $P(A|B) = P(A)$.

A definição de probabilidade condicional possui características interessantes, primeiramente o fato da $P(A \cap B) = P(A)$ quando $P(B) = 0$ permite que a probabilidade condicional tenha semelhanças com a notação usual de uma função uma vez que a probabilidade condicional $P(A \cap B)$ dependerá somente de A [5]. Outro fator de relevância deve-se ao interesse teórico uma vez que probabilidade condicional permitir a decomposição de probabilidades que possuem uma difícil caracterização por meio de probabilidade condicionais mais simples[6]. Por fim, baseando-se na probabilidade condicional é possível o produto de probabilidades.

Proposição 3.0.1 (Regra do produto de probabilidades [6]). Para os eventos A_1, A_2, \dots, A_n em (Ω, \mathcal{F}, P) com $P(\bigcap_{i=1}^{\infty}) > 0$. O produto das probabilidades é dado por,

$$P(A_1 \cap A_2 \cap A_3 \cdots \cap A_n) = P(A_1)P(A_2|A_1) \dots P(A_n|A_1 \cap A_2 \dots A_{n-1}).$$

Uma aplicação direta da regra do produto de probabilidades é a lei da probabilidade total.

Teorema 3.0.1 (Lei da Probabilidade Total[6]). Supondo-se que os eventos C_1, C_2, \dots, C_n em (Ω, \mathcal{F}, P) formam uma partição de Ω e que para qualquer C_n tem-se $C_n > 0$. Então, para qualquer evento A neste espaço de probabilidade. A probabilidade do

evento A é dada por,

$$P(A) = \sum_{i=1}^n P(C_i)P(A|C_i).$$

O membro do direito lei de probabilidade total é formada produtos envolvendo as probabilidades condicionais. Por meio da probabilidade condicional pode-se caracterizar de outra forma a independência de eventos.

Definição 3.0.5 (Independência de dois eventos[6]). Dados dois eventos A e B em (Ω, F, P) são ditos independentes quando a ocorrência do evento A não influencia na ocorrência do evento B, isto é,

$$P(A|B) = P(A). \quad (3.3)$$

Conforme a definição 3.0.4 de probabilidade condicional,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (3.4)$$

de 3.3 e 3.4 obtêm-se uma representação equivalente para a independência de eventos dada por,

$$P(A \cap B) = P(A)P(B).$$

Portanto, a independência para dois eventos pode ser analisada utilizando-se o cálculo da probabilidade associada com a intersecção facilitando-se a análise dos eventos e futuras conclusões sobre o fenômeno observado. Com base na definição 3.0.5 é definida a independência de vários eventos.

Definição 3.0.6. Os eventos A_1, A_2, \dots, A_n em (Ω, F, P) são independentes se para toda coleção de índices $1 \leq i_1 \leq i_2 < \dots i_k \leq n$ é verdadeiro que

$$P(A_{i_1} \cap A_{i_2} \cap A_{i_3} \dots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k}).$$

Consequentemente por 3.0.5 e 3.0.6, é possível analisar a independência de eventos com base no cálculo de probabilidades. Outro fator importante deve-se que pelo conhecimento da independência eventos facilitar a análise do evento de interesse, por exemplo, uma moeda é lançada duas vezes. Sejam os eventos:

A : Sair cara;

B : Sair coroa.

Nota-

se que os eventos A e B são independentes. De fato, a ocorrência do evento A ou B no primeiro lançamento em nada influencia o resultado para o segundo lançamento. A probabilidade de sair cara no segundo lançamento dado que saiu coroa no primeiro

lançamento corresponde a $\frac{1}{4}$. Pois, como os eventos são indenpendentes,

$$\begin{aligned} P(B|A) &= P(B) \\ P(B|A) &= \frac{1}{2}. \end{aligned}$$

é de imediato que,

$$\begin{aligned} P(B \cap A) &= P(B)P(A) \\ P(B \cap A) &= \left(\frac{1}{2}\right) \left(\frac{1}{2}\right) = \frac{1}{4} \end{aligned}$$

O mesmo resultado poderia ser obtido calculando-se a probabilidade pela forma clássica, ou seja elecando-se os resultados possíveis, (cara, cara), (cara, coroa), (coroa, cara), (coroa, coroa). O evento de interesse é (coroa, cara) para o experimento $\Omega = 4$ tem-se de imediato que,

$$P(\text{coroa, cara}) = \frac{1}{4}.$$

A primeira forma e a segunda tiveram o mesmo resultado, contudo, se o evento em um espaço amostral é constituído por quantidade de elementos suficientemente grande a primeira forma é mais atraente. No estudo de fenômenos aletórios é comum o interesse em quantidades que possam ser associadas ao evento. Antes da realização de qualquer experimento ou ocorrência de qualquer fenômeno de natureza aleatória não é comum ter o resultado. Contudo, uma vez que o espaço de probabilidade esteja bem definido é possível observar qualquer evento de interesse na σ -álgebra. Dessa forma, também é possível atribuir probabilidade para as funções que descrevem o evento. Nessa perspectiva fundamenta-se o conceito do termo variável aleatória.

Definição 3.0.7 (Variável aleatória). Dado um espaço de probabilidade (Ω, F, P) , define-se uma variável aleatória como uma função $X: \Omega \rightarrow \mathbb{R}$ tal que

$$X^{-1}(I) = \{\omega \in \Omega; X(\omega) \in I\} \in F.$$

qualquer que seja o conjunto $I \subset \mathbb{R}$ pertecente a σ -álgebra.

Pela definição 3.0.7 existe uma função de Ω em \mathbb{R} . Pela qual é possível ser estabelecer sua inversa de Ω em \mathbb{R} . A exigência que a inversa pertence a σ -álgebra é importante, pois é possível garantir as operações com as probabilidades para os elementos que estejam na σ -álgebra. Considerando-se $\Omega = \{1, 2, 3, 4\}$ e seja $F = \{\emptyset, \Omega, \{1, 2\}, \{3, 4\}\}$ definindo-se os conjuntos $A = \{1, 2\}$ e $B = \{1, 3\}$ deve-se verificar se é possível I_A e I_B sejam variáveis aleatórias[6]. Primeiramente,

definindo-se X^{-1} por

$$(\omega \in \Omega; I_A \in (-\infty, x]) = \begin{cases} \emptyset, & \text{se } x < 0, \\ A^c, & \text{se } 0 \leq x < 1, \\ \Omega, & \text{se } x \geq 1 \end{cases}$$

nota-se que para qualquer intervalo da forma $(-\infty, x] \in F$ uma vez que ambos os subconjuntos $\emptyset, A^c, \Omega \in F$ estão em F , dessa forma I_A é variável aleatória. Contudo, utilizando-se a mesma definição de I_A em I_B ocorrer que B^c não pertence a F . De fato, $B^c = \{2, 4\} \notin F$ nessas condições I_B não é uma variável aleatória. Por outro lado, poderia-se definir uma outra σ -álgebra que comporta-se o I_B permitindo-se que I_B seja uma variável aleatória. A formulação de uma variável aleatória está sempre associada diretamente com a σ -álgebra. Os eventos da forma $(-\infty, x)$ possuem um interesse em particular, pois, constituem uma importante função no estudo da probabilidade denominada função de distribuição de probabilidade. No intuito de facilitar o entendimento quando houve menção a ω em algum evento da forma $\{\omega \in \Omega; X(\omega) \in I\}$ este será omitido, portanto a representação do evento de interesse será $[X \in I]$. Desta forma, os eventos como $I = (-\infty, x]$ seram representados por $[X < 0]$ em vez de $P(\{\omega \in \Omega : X(\omega) \in I\})$ será adotada a nomenclatura $P(X \in I)$. Por este tipo de simplificação a função de distribuição de probabilidade também é conhecida por função de probabilidade acumulada pelo acúmulo de probabilidade no intervalo real[6].

Definição 3.0.8 (Função de distribuição de probabilidade). Seja X uma variável aleatória em um espaço de probabilidade (Ω, F, P) . A função de distribuição é dada por,

$$F_x(x) = P(X \in (-\infty, x]) = P(X \leq x).$$

O conhecimento da definição 3.0.8 permite obter qualquer informação sobre a variável a ser examinada. Apesar da variável aleatória assumir valores num subconjunto dos reais a sua função de distribuição assumir valores em todos os reais [6]. Além da função de distribuição existe uma outra função de importância semelhante para o estudo da probabilidade conhecida como função de probabilidade. Contudo, deve ser ressaltado que uma variável aleatória X é classificada como: discreta ou contínua. Uma variável é classificada como discreta quando assumir somente um número enumerável de valores finito ou infinito. Desta forma, uma função de probabilidade para esse tipo de variável aleatória atribuir valor a cada um dos possíveis valores assumidos pela variável aleatória, isto é, tomando-se X como uma variável aleatória que assumir valores em x_1, x_2, \dots , para $i = 1, 2, \dots$, portanto, a função de

probabilidade é dada por,

$$P(x_i) = P(X = x_i) = P(\omega \in \Omega; X(\omega) = x_i).$$

Em relação ao caso discreto a função de probabilidade deve satisfazer duas propriedades,

$$p_1: 0 \leq p(x_i) \leq 1, \forall i = 1, 2, \dots,$$

$$p_2: \sum_i p(x_i) = 1.$$

As propriedades p_1 e p_2 confirmam os axiomas de probabilidade. De fato, por p_1 a probabilidade somente assumirá valores não negativos no intervalo $[0, 1]$ e por p_2 o somatório de todas as probabilidades associadas ao evento deve ser igual a 1. Para o caso contínuo a ideia é análoga, contudo como a variável aleatória é contínua a utilização de uma integral ao invés do somatório é o ideal, portanto, uma variável aleatória X será classificada como contínua quando existe uma função f não negativa que satisfaça,

$$p_1 : \int_{-\infty}^{\infty} f(\omega) d\omega \geq 0, \forall x \in \mathbb{R},$$

$$p_2 : \int_{-\infty}^{\infty} f(\omega) d\omega = 1.$$

Observando-se que a construção de tais funções deve existir um espaço de probabilidade (Ω, \mathcal{F}, P) bem definido. Pelas funções de distribuição e de probabilidade consegue-se a maioria das informações sobre o evento. Por fim, com fundamentação adquirida pode-se ser construído um dos principais conceitos da teoria da probabilidade o de valor esperado, esperança matemática, ou média de uma variável extremamente útil na análise de fenômenos aleatórios justamente por se comportar como a média permitindo informações do problema em um contexto geral. Podendo ser definido para o caso discreto como,

$$E(X) = \sum_{i=1}^{\infty} x_i p_X(x_i),$$

desde que soma esteja bem determinada e o espaço de probabilidade bem definido. Para o caso contínuo, a ideia é análoga,

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx,$$

desde que a integral esteja bem definida, assim como o espaço de probabilidade. Nesse capítulo foram abordados os principais conceitos de probabilidade relevantes

para pesquisa. Primeiramente, a construção da ideia de probabilidade bem como as relações que fundamentam seu uso. Além, de estrutura os conceitos como o de valor de esperado de fundamental importância para o despacho de energia no cálculo de custo esperado.

Referências Bibliográficas

- [1] AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA, 2008, *Atlas de Energia Elétrica*, 3 ed. Brasília.
- [2] BOYD, S. P., 2009, *Convex Optimization*. New York, Cambridge University.
- [3] CENTRO DE PESQUISAS DE ENERGIA ELÉTRICA, 2001, *Modelo DE-COMP CEPEL:Manual de referência*.
- [4] IZMAILOV, A., SOLOV, M., 2014, *Otimização volume 1. Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. 3 ed. Rio de janeiro, IMPA.
- [5] JAMES, B. R., 1981, *Probabilidade: um curso em nível intermediário*. 4 ed. Rio de janeiro, IMPA.
- [6] MAGALHÃES, M. N., 2013, *Probabilidade e Variáveis Aleatórias*. 3 ed. São Paulo, Editora da Universidade de São Paulo.
- [7] TOLMASQUIM, M., 2016, *Energia Termelétrica:Gás Natural, Biomassa, Carvão, Nuclear*. Rio de janeiro, EPE.